

SCHWACH DOMINIERTE FAMILIEN VON MASSEN

Dieter Mussmann

1971



Schwach dominierte Familien von Maßen

Inaugural-Dissertation

zur Erlangung der Doktorwürde der

Naturwissenschaftlichen Fakultät

der Ludwig-Maximilians-Universität München



vorgelegt von

Dieter Mußmann

München, Januar 1971



Berichterstatter: Priv. - Doz. Dr. V. Mammitsch

Prof. Dr. K. Jörgens

Mündliche Prüfung: 3.2. 1971

## Inhaltsverzeichnis

|                                                                                                                | Seite |
|----------------------------------------------------------------------------------------------------------------|-------|
| Problemstellung und Ergebnisse                                                                                 | 1     |
| Bezeichnungen                                                                                                  | 5     |
| I. Suffiziente Experimente                                                                                     | 8     |
| § 1. Einbettung eines Experimentes in ein schwach dominiertes Experiment                                       | 8     |
| § 2. Eigenschaften suffizienter Unter- $\sigma$ -Algebren für schwach dominierte Familien von normierten Maßen | 17    |
| § 3. Suffizienz und paarweise Suffizienz                                                                       | 23    |
| § 4. Ein verallgemeinertes Neyman-Kriterium                                                                    | 27    |
| § 5. Minimalsuffizienz                                                                                         | 33    |
| II. Schwach suffiziente Experimente                                                                            | 38    |
| § 6. Schwache Suffizienz beim Vergleich zweier Experimente                                                     | 38    |
| § 7. Schwache Suffizienz und paarweise schwache Suffizienz                                                     | 51    |
| § 8. $f$ -Suffizienz                                                                                           | 53    |
| § 9. Ein Isomorphiesatz                                                                                        | 57    |
| III. Entscheidungstheoretische Charakterisierung der schwachen Suffizienz                                      | 63    |
| § 10. Hilfssätze über spezielle lineare und bilineare Abbildungen                                              | 63    |
| § 11. Entscheidungsfunktionen                                                                                  | 69    |
| § 12. Entscheidungstheoretische Charakterisierung der schwachen Suffizienz                                     | 71    |
| Literaturverzeichnis                                                                                           | 74    |



## Problemstellung und Ergebnisse

Bei der mathematischen Beschreibung naturwissenschaftlicher Experimente benötigt man häufig einen Meßraum  $(X, \mathcal{A})$ , auf dem eine nicht leere Familie von normierten Maßen  $(P_\beta)_{\beta \in \Theta}$  gegeben ist. In Anlehnung an Blackwell ([3],[4]) nennen wir das Tripel  $(X, \mathcal{A}, (P_\beta)_{\beta \in \Theta})$  selbst ein Experiment, ohne uns dabei um Anwendungsmöglichkeiten zu kümmern. Wenn  $\mathcal{G}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$  ist, bezeichnen wir  $(X, \mathcal{G}, (P_\beta|_{\mathcal{G}})_{\beta \in \Theta})$  als Unterexperiment.  $\mathcal{G}$  heißt bekanntlich suffizient für  $(P_\beta)_{\beta \in \Theta}$ , wenn es für jede  $\mathcal{A}$ -meßbare Indikatorfunktion eine  $\mathcal{G}$ -meßbare Funktion gibt, die für jene bedingter Erwartungswert bezüglich  $\mathcal{G}$  für jedes  $P_\beta$ ,  $\beta \in \Theta$ , ist. In diesem Fall reden wir auch von einem suffizienten Unterexperiment. Wenn  $(P_\beta)_{\beta \in \Theta}$  dominiert ist, d.h. wenn jedes  $P_\beta$ ,  $\beta \in \Theta$ , eine Dichte bezüglich eines festen  $\sigma$ -endlichen Maßes auf  $(X, \mathcal{A})$  hat, gibt es einige Sätze über für  $(P_\beta)_{\beta \in \Theta}$  suffiziente Unter- $\sigma$ -Algebren ([1],[7]), die in allgemeineren Fällen nicht ohne weiteres gelten ([5],[7],[13]).

Wir stellen uns das Problem, inwieweit diese Sätze an die  $\sigma$ -Endlichkeit des dominierenden Maßes gebunden sind. Eine Verallgemeinerung der  $\sigma$ -endlichen Maße sind die lokalisierbaren Maße mit der endlichen Teilmengeneigenschaft ([15], § 34 und § 35). Wir nennen nun die Familie  $(P_\beta)_{\beta \in \Theta}$  und auch das Experiment selbst schwach dominiert, wenn es auf  $(X, \mathcal{A})$  ein lokalisierbares Maß mit der endlichen Teilmengeneigenschaft gibt, bezüglich dessen jedes  $P_\beta$ ,  $\beta \in \Theta$ , eine Dichte hat. Eine für eine schwach dominierte Familie  $(P_\beta)_{\beta \in \Theta}$  suffiziente Unter- $\sigma$ -Algebra  $\mathcal{G}$  hat dann die Eigenschaft, daß

$(P_{\mathfrak{A}}|\mathfrak{K})_{\mathfrak{A} \in \mathcal{O}}$  ebenfalls schwach dominiert ist (2.4), und wenn eine beliebige Unter- $\sigma$ -Algebra diese Eigenschaft hat, können wir zeigen, daß ihre paarweise Suffizienz für  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$ , d.h. ihre Suffizienz für jede zweielementige Teilfamilie von  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$ , ihre Suffizienz für  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$  liefert (3.5) und daß diese auch aus einer gewissen Faktorisierbarkeit der Dichten der  $P_{\mathfrak{A}}$ ,  $\mathfrak{A} \in \mathcal{O}$ , folgt (4.5). Weiter läßt sich für eine schwach dominierte Familie  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$  die Existenz einer minimalsuffizienten  $\sigma$ -Algebra zeigen (5.3), die durch die Meßbarkeit gewisser Dichten charakterisiert werden kann (5.4). Dieses sind Verallgemeinerungen der für dominierte Familien bekannten Ergebnisse.

Ein Experiment ist mit einem suffizienten Unterexperiment vermöge eines bedingten Erwartungswertes verknüpft. Allgemeiner kann man zwei Experimente mit gleicher Indexmenge betrachten, die in analoger Weise miteinander verknüpft sind. Wir schließen uns hierbei Morse und Sacksteder [12] an, die Experimente untersuchen, die durch schwache Markovkerne (6.1) miteinander verbunden sind. Es scheinen zwei derartige Verknüpfungen sinnvoll. In einem Fall werden die Experimente isomorph genannt (6.4), und im anderen Fall wird das eine Experiment für das andere schwach suffizient genannt (6.8)(1). Dem Begriff der paarweisen Suffizienz entspricht der Begriff der paarweisen schwachen Suffizienz (6.8)(2). (In den Benennungen sind wir von Morse und Sacksteder abgewichen.) Sacksteder [14] zeigt, daß bei Dominiertheit der auftretenden Familien von normierten Maßen paarweise schwache Suffizienz, Isomorphie und schwache Suffi-

zienz äquivalent sind. Wir erhalten dies Ergebnis auch für den Fall, daß ein Experiment schwach dominiert ist (7.1), (7.2). Jedes suffiziente Unterexperiment ist für das Ausgangsexperiment auch schwach suffizient, und bei einem Experiment, das schwach dominiert ist, enthält jedes dafür schwach suffiziente Unterexperiment bis auf Nullmengen ein suffizientes Unterexperiment (6.11). Paarweise Suffizienz und auch paarweise schwache Suffizienz lassen sich nach Csiszár [6] informationstheoretisch charakterisieren. Die Äquivalenz von paarweiser schwacher Suffizienz und schwacher Suffizienz im schwach dominierten Fall liefert so eine informationstheoretische Charakterisierung der schwachen Suffizienz (8.8), die für dominierte Familien bekannt ist [8]. Für den Fall, daß zwei Experimente füreinander schwach suffizient und beide schwach dominiert sind, zeigen wir, daß die minimalsuffizienten Unterexperimente maßisomorph sind (9.3). Schließlich kommen wir durch Umformung eines Satzes von Heyer [8] zu dem Ergebnis, daß für zwei Experimente, die schwach dominiert sind, die schwache Suffizienz vermöge eines Vergleichs von "Risikofunktionen" der Experimente bei einer gewissen Klasse von "Entscheidungsräumen" charakterisiert werden kann (12.6).

Zu erwähnen ist noch die folgende Beziehung zur Arbeit von Le Cam [11]. Le Cam betrachtet als Experiment ein Tripel  $(X, E, (P_\theta)_{\theta \in \Theta})$ , welches aus einer nicht leeren Menge  $X$ , einem Banachverband  $E$  beschränkter reeller Funktionen auf  $X$  mit 1 und einer nicht leeren Familie  $(P_\theta)_{\theta \in \Theta}$  von positi-

ven, normierten Linearformen auf  $E$  besteht. Diesem Experiment ordnet er zwei Banachverbände  $L$  und  $M$  zu mit  $L' = M$  derart, daß  $(P_j)_{j \in \mathbb{Q}}$  eine Teilmenge von  $L$  bildet und jedes Element von  $E$  ein Element aus  $M$  liefert. Experimente untersucht er dann auf Grund dieser Räume. Grob gesagt heißt das, er ersetzt die Untersuchung beliebiger Experimente durch die Untersuchung dieses Raumpaars. Bezüglich dieser Einbettung eines Experimentes in ein Raumpaar  $L, M$  zeigen wir, daß es (auf rein maßtheoretischem Wege) möglich ist, ein in unserem Sinne definiertes Experiment in ein schwach dominiertes Experiment einzubetten (1.10). Der  $L^1$ - bzw.  $L^\infty$ -Raum bezüglich eines Maßes, das die Familie von normierten Maßen dieses Experimentes schwach dominiert und dazu äquivalent ist, entspricht dann dem Raum  $L$  bzw.  $M$  bei  $Le\ Cam$ .

## Bezeichnungen

$(X, \mathcal{O})$  sei im folgenden immer ein Meßraum,  $\mathcal{O}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{O}$  und  $(\nu_i)_{i \in I}$ ,  $I \neq \emptyset$ , und  $(\mu_j)_{j \in J}$ ,  $J \neq \emptyset$ , seien zwei Familien von Maßen auf  $(X, \mathcal{O})$ . Wenn die Indexmengen einelementig sind, schreiben wir  $\nu$  bzw.  $\mu$ .

Eine Menge  $N \in \mathcal{O}$  heißt  $(\nu_i)_{i \in I}$ -Nullmenge, wenn  $\nu_i(N) = 0$  ist für alle  $i \in I$ . Gilt eine Eigenschaft für alle  $x \in X$  bis auf eine  $(\nu_i)_{i \in I}$ -Nullmenge, so sagen wir, sie gilt  $(\nu_i)_{i \in I}$ -fast überall.  $(\nu_i)_{i \in I}$  heißt absolut stetig bezüglich  $(\mu_j)_{j \in J}$ , wenn jede  $(\mu_j)_{j \in J}$ -Nullmenge auch eine  $(\nu_i)_{i \in I}$ -Nullmenge ist.  $(\nu_i)_{i \in I}$  heißt äquivalent zu  $(\mu_j)_{j \in J}$ , wenn die  $(\nu_i)_{i \in I}$ -Nullmengen genau die  $(\mu_j)_{j \in J}$ -Nullmengen sind.

Eine  $\mathcal{O}$ -meßbare Funktion, die für alle  $i \in I$  mit  $\nu_i$  integriert werden kann, wird  $(\nu_i)_{i \in I}$ -integrierbar genannt.  $(\nu_i)_{i \in I}$  sei nun eine Familie von endlichen Maßen auf  $(X, \mathcal{O})$  und  $f$  eine  $(\nu_i)_{i \in I}$ -integrierbare Funktion. Eine  $\mathcal{O}$ -meßbare Funktion, die wir mit  $E(f|\mathcal{O})$  bezeichnen, heißt bedingter  $(\nu_i)_{i \in I}$ -Erwartungswert von  $f$  bezüglich  $\mathcal{O}$ , wenn für alle  $S \in \mathcal{O}$  und  $i \in I$  gilt  $\int_S E(f|\mathcal{O}) d\nu_i = \int_S f d\nu_i$ .  $E(f|\mathcal{O})$  ist, sofern es überhaupt existiert, im allgemeinen nicht eindeutig bestimmt.  $E(f|\mathcal{O})$  soll daher immer eine feste Version der möglichen Funktionen bedeuten.

$\mathcal{O}$  heißt suffizient für  $(\nu_i)_{i \in I}$ , wenn für alle  $A \in \mathcal{O}$  ein bedingter  $(\nu_i)_{i \in I}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{O}$  existiert. Mit  $1_A$  bezeichnen wir dabei die Indikatorfunktion von  $A \in \mathcal{O}$ .

Gilt für  $E, F \in \mathcal{O}$   $\nu_i((E \cap F) \Delta E) = 0$  bzw.  $\nu_i(E \Delta F) = 0$

für alle  $i \in I$ , so schreiben wir  $E \subset F (v_i)_{i \in I}$ -fast bzw.  $E = F (v_i)_{i \in I}$ -fast.  $\mathcal{E}$  und  $\mathcal{F}$  seien zwei Teilmengen von  $\mathcal{A}$ . Wenn es zu einem  $A \in \mathcal{A}$  ein  $F \in \mathcal{F}$  mit  $A = F (v_i)_{i \in I}$ -fast gibt, so sagen wir  $A \in \mathcal{F} (v_i)_{i \in I}$ -fast. Falls für jedes  $E \in \mathcal{E}$  gilt  $E \in \mathcal{F} (v_i)_{i \in I}$ -fast, benutzen wir die Schreibweise  $\mathcal{E} \subset \mathcal{F} (v_i)_{i \in I}$ -fast. Wenn  $\mathcal{E} \subset \mathcal{F} (v_i)_{i \in I}$ -fast und  $\mathcal{F} \subset \mathcal{E} (v_i)_{i \in I}$ -fast ist, schreiben wir  $\mathcal{E} = \mathcal{F} (v_i)_{i \in I}$ -fast.

Sind  $\nu$  und  $\mu$  zwei Maße auf  $(X, \mathcal{A})$  und hat  $\nu$  eine Dichte bezüglich  $\mu$ , so bezeichnet  $\frac{d\nu}{d\mu}$  stets eine Festlegung dieser Dichte.

Vermöge eines Maßes  $\nu$  auf  $(X, \mathcal{A})$  erhält man bekanntlich eine Einteilung der  $\mathcal{A}$ -meßbaren Funktionen und auch von  $\mathcal{A}$  selbst in Äquivalenzklassen. Wir bezeichnen für jede  $\mathcal{A}$ -meßbare Funktion  $f$  mit  $\tilde{f}^\nu$  die Menge der  $\mathcal{A}$ -meßbaren Funktionen, die mit  $f$   $\nu$ -fast überall übereinstimmen. Wir sagen,  $\tilde{f}^\nu$  ist die  $\nu$ -Äquivalenzklasse von  $f$ . Wenn klar ist, bezüglich welchen Maßes die Äquivalenzklassen gebildet werden, lassen wir das  $\nu$  weg. Die Menge der  $\nu$ -Äquivalenzklassen  $\mathcal{A}$ -meßbarer Funktionen hat eine teilweise Ordnung  $\leq$ , die durch  $\tilde{f} \leq \tilde{g}$ , wenn  $f \leq g$   $\nu$ -fast überall ist, definiert ist.

Mit Hilfe der Indikatorfunktionen überträgt sich diese teilweise Ordnung auf  $\nu$ -Äquivalenzklassen von Elementen aus  $\mathcal{A}$ . Mit  $\tilde{A}^\nu$ ,  $A \in \mathcal{A}$ , bezeichnen wir die Menge aller  $B \in \mathcal{A}$  mit  $\tilde{1}_A^\nu = \tilde{1}_B^\nu$ , und es sei  $\tilde{A}^\nu \leq \tilde{B}^\nu$ , wenn  $\tilde{1}_A^\nu \leq \tilde{1}_B^\nu$  ist.

Suprema und Infima von Familien von  $\nu$ -Äquivalenzklassen  $\mathcal{A}$ -meßbarer Funktionen oder von  $\nu$ -Äquivalenzklassen von Elementen aus  $\mathcal{A}$  sind immer bezüglich der obigen teilweisen Ordnungen zu verstehen.

Ein Maß  $\nu$  auf  $(X, \mathcal{A})$  sei gegeben. Wir nennen  $\nu$  lokalisierbar, wenn für jede Familie  $(A_i)_{i \in I}$ ,  $I \neq \emptyset$ , von Mengen aus  $\mathcal{A}$   $\sup_{i \in I} \widetilde{A}_i$  existiert.

Man sagt,  $\nu$  hat die endliche Teilmengeneigenschaft, falls zu jedem  $A \in \mathcal{A}$  mit  $\nu(A) > 0$  eine Teilmenge  $B$  von  $A$ ,  $B \in \mathcal{A}$ , existiert mit  $\infty > \nu(B) > 0$ . (In diesem Fall ist unsere Definition von lokalisierbar mit der in [15] äquivalent.)

Für ein festes  $A \in \mathcal{A}$  sei  $\nu^A$  das durch  $\nu^A(B) := \nu(A \cap B)$  für alle  $B \in \mathcal{A}$  definierte Maß.

Wir sagen,  $A \in \mathcal{A}$  hat  $\sigma$ -endliches  $\nu$ -Maß, wenn  $A$  eine abzählbare disjunkte Vereinigung von Mengen aus  $\mathcal{A}$  mit endlichen  $\nu$ -Mäßen ist.

Ist  $(\nu_i)_{i \in I}$ ,  $I \neq \emptyset$ , eine Familie von endlichen Mäßen auf  $(X, \mathcal{A})$ , so erhält man auf  $\mathcal{A}$  eine Äquivalenzrelation, wenn man Mengen  $A, B \in \mathcal{A}$  äquivalent nennt, falls  $A = B \setminus (\nu_i)_{i \in I}$ -fast ist. Die so entstehenden Äquivalenzklassen bezeichnen wir als  $(\nu_i)_{i \in I}$ -Äquivalenzklassen, und für deren Gesamtheit verwenden wir das Symbol  $\mathcal{A}((\nu_i)_{i \in I})$ .  $\widetilde{A}$  sei die  $(\nu_i)_{i \in I}$ -Äquivalenzklasse von  $A \in \mathcal{A}$ . Für  $\widetilde{A} \in \mathcal{A}((\nu_i)_{i \in I})$  sei  $\widetilde{X} \setminus \widetilde{A} := \widetilde{X} \setminus A$  und  $\nu_i(\widetilde{A}) := \nu_i(A)$  für alle  $i \in I$ . Für  $A_n \in \mathcal{A}((\nu_i)_{i \in I})$ ,  $n = 1, 2, \dots$ , sei  $\bigvee_{n=1}^{\infty} \widetilde{A}_n := \bigcup_{n=1}^{\infty} A_n$ .

Bezüglich der Definition von Integralen verweisen wir auf Bauer [2].

## I. Suffiziente Experimente

### § 1. Einbettung eines Experimentes in ein schwach dominiertes Experiment

Im Zusammenhang mit naturwissenschaftlichen Experimenten ist es häufig notwendig, auf einem vorgegebenen Meßraum mehrere normierte Maße gleichzeitig zu betrachten. Um dafür einen kurzen Namen zu haben, führen wir den Begriff "Experiment" ein, ohne allerdings auf seine Interpretation und Anwendungsmöglichkeiten in den Naturwissenschaften näher einzugehen.

(1.1) Definition. (a) Ein Tripel  $(X, \mathcal{A}, (P_\lambda)_{\lambda \in \mathcal{Q}})$  heißt Experiment, wenn  $(X, \mathcal{A})$  ein Meßraum ist und  $(P_\lambda)_{\lambda \in \mathcal{Q}}$ ,  $\mathcal{Q} \neq \emptyset$ , eine Familie darauf definierter normierter Maße. (b) Ist  $\mathcal{G}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ , so heißt das Tripel  $(X, \mathcal{G}, (P_\lambda|_{\mathcal{G}})_{\lambda \in \mathcal{Q}})$  ein Unterexperiment von  $(X, \mathcal{A}, (P_\lambda)_{\lambda \in \mathcal{Q}})$ .

(1.2) Definition. Das Unterexperiment heißt suffizient, wenn  $\mathcal{G}$  suffizient für die Familie  $(P_\lambda)_{\lambda \in \mathcal{Q}}$  ist.

(1.3) Definition. Eine Familie von endlichen Maßen  $(\nu_i)_{i \in I}$ ,  $I \neq \emptyset$ , auf einem Meßraum  $(X, \mathcal{A})$  heißt schwach dominiert, wenn es auf  $\mathcal{A}$  ein lokalisierbares Maß  $\nu$  mit der endlichen Teilmengeneigenschaft gibt, bezüglich dessen jedes  $\nu_i$ ,  $i \in I$ , eine Dichte hat.  $\nu$  heißt schwach dominierendes Maß für  $(\nu_i)_{i \in I}$ .

(1.4) Definition. Ein Experiment  $(X, \mathcal{A}, (P_\lambda)_{\lambda \in \mathcal{Q}})$  heißt schwach dominiert, wenn die Familie  $(P_\lambda)_{\lambda \in \mathcal{Q}}$  schwach dominiert ist.

$(X, \mathcal{A}, (P_\lambda)_{\lambda \in \mathcal{Q}})$  sei ein Experiment. Um dies in ein schwach

dominiertes Experiment einzubetten, benötigen wir einige Eigenschaften der teilweise geordneten Menge  $\mathfrak{M}$  aller endlichen Maße auf  $(X, \mathcal{A})$ , zu der ja jedes  $P_\mu$ ,  $\mu \in \mathcal{Q}$ , gehört. Die teilweise Ordnung  $\leq$  von  $\mathfrak{M}$  ist dabei definiert durch  $P \leq Q$ , bei  $P, Q \in \mathfrak{M}$ , wenn  $P(A) \leq Q(A)$  für alle  $A \in \mathcal{A}$  ist.

Man kann diese teilweise Ordnung mit Hilfe von Dichten beschreiben.  $P, Q \in \mathfrak{M}$  seien bezüglich des endlichen Maßes  $\mu$  absolut stetig (ein solches  $\mu$  existiert stets). Dann ist  $P \leq Q$  gleichbedeutend mit  $\frac{dP}{d\mu} \leq \frac{dQ}{d\mu}$   $\mu$ -fast überall.

Weiter existiert das Infimum zweier Maße  $P, Q \in \mathfrak{M}$ , und es ist  $\frac{d(\inf(P, Q))}{d\mu} = \inf(\frac{dP}{d\mu}, \frac{dQ}{d\mu})$   $\mu$ -fast überall.

$P, Q \in \mathfrak{M}$  heißen orthogonal, wenn  $\inf(P, Q)$  das Nullmaß ist.

(1.5) Bemerkung. Die Glieder einer Folge  $(Q_n)_{n \in \mathbb{N}}$  von Maßen aus  $\mathfrak{M}$  sind genau dann paarweise orthogonal, wenn es eine Zerlegung  $(A_m)_{m \in \mathbb{N}}$ ,  $A_m \in \mathcal{A}$ , von  $X$  gibt so, daß  $Q_n(A_m) = 0$  für  $n \neq m$  ist.

Dies kann man sich durch die Betrachtung von Dichten der  $Q_n$ ,  $n \in \mathbb{N}$ , bezüglich eines passenden endlichen Maßes, zum Beispiel  $\sum_{n=1}^{\infty} c_n Q_n$  mit  $\sum_{n=1}^{\infty} c_n Q_n(X) < \infty$ ,  $c_n > 0$  für alle  $n \in \mathbb{N}$ , überlegen.

Zu  $P, Q \in \mathfrak{M}$  kann man ein Maß  $Pr_P(Q) \in \mathfrak{M}$  definieren, das grob gesagt eine Projektion von  $Q$  auf  $P$  darstellt (Kakutani [10] behandelt diese Konstruktion in einem allgemeineren Rahmen).

(1.6) Definition.  $P, Q \in \mathfrak{M}$  seien bezüglich  $\mu \in \mathfrak{M}$  absolut stetig. Für alle  $A \in \mathcal{A}$  wird gesetzt:

$$Pr_P(Q)(A) := \lim_{n \rightarrow \infty} \int_A \inf(n \frac{dP}{d\mu}(x), \frac{dQ}{d\mu}(x)) \mu(dx).$$

$Pr_P(Q)$  heißt Projektion von  $Q$  auf  $P$ .

Die Definition hängt nicht von  $\mu$  ab, und man hat überdies:

(1.7) Folgerung. (a) Es ist stets  $\text{Pr}_P(Q) \leq Q$ .

(b)  $\text{Pr}_P(Q) = Q$  gilt genau dann, wenn  $Q$  bezüglich  $P$  absolut stetig ist.

(1.8) Lemma. Sind  $P_1, P_2 \in \mathfrak{M}$  orthogonal, so sind auch  $\text{Pr}_{P_1}(Q)$  und  $\text{Pr}_{P_2}(Q)$  orthogonal,  $Q \in \mathfrak{M}$ .

Beweis. Gemäß (1.5) gibt es ein  $A \in \mathfrak{A}$  so, daß  $P_1(X \setminus A) = 0$  und  $P_2(A) = 0$  ist. Aus (1.7)(a) schließt man dann  $\text{Pr}_{P_1}(Q)(X \setminus A) = 0$  und  $\text{Pr}_{P_2}(Q)(A) = 0$ , und daraus ergibt sich die Orthogonalität dieser Maße.

Wir können jetzt eine zu  $(P_\delta)_{\delta \in \Theta}$  äquivalente Familie von endlichen Maßen auf  $(X, \mathfrak{A})$  bestimmen, die für die Einbettung von  $(X, \mathfrak{A}, (P_\delta)_{\delta \in \Theta})$  in ein schwach dominiertes Experiment wichtig sein wird.

(1.9) Satz.  $(X, \mathfrak{A}, (P_\delta)_{\delta \in \Theta})$  sei ein Experiment und  $\mathcal{J}$  eine suffiziente Unter- $\sigma$ -Algebra von  $\mathfrak{A}$  für  $(P_\delta)_{\delta \in \Theta}$ . Dann existieren eine Teilfamilie  $(P_{\delta_i})_{i \in I}$ ,  $I \neq \emptyset$ , von  $(P_\delta)_{\delta \in \Theta}$  und eine Familie  $(S_i)_{i \in I}$ ,  $S_i \in \mathcal{J}$ , derart, daß  $(P_{\delta_i}^{S_i})_{i \in I}$  paarweise orthogonal,  $P_{\delta_i}^{S_i}$  für alle  $i \in I$  vom Nullmaß verschieden und jedes  $P_\delta$ ,  $\delta \in \Theta$ , bezüglich einer höchstens abzählbaren Teilfamilie von  $(P_{\delta_i}^{S_i})_{i \in I}$  absolut stetig ist.

Beweis. Wir betrachten die Menge  $Z$  aller paarweise orthogonalen Familien von endlichen positiven Maßen der Gestalt  $(P_{\delta_j}^{S_j})_{j \in J}$ ,  $J \neq \emptyset$ ,  $S_j \in \mathcal{J}$ . Familien, die aufgefaßt als Mengen übereinstimmen, seien dabei nicht unterschieden.

$Z$  ist nicht leer, denn für ein  $\delta_0 \in \Theta$  ist  $P_{\delta_0}^X = P_{\delta_0}$  in  $Z$ .

Auf  $Z$  wird eine teilweise Ordnung  $\leq$  definiert durch

$(P_{\delta_j}^{S_j})_{j \in J} \leq (P_{\delta_k}^{S_k})_{k \in K}$ , wenn zu jedem  $j_0 \in J$  ein  $k_0 \in K$  exi-

stiert mit  $P_{j_0}^{S_{j_0}} = P_{j_0}^{S_{k_0}}$ . Jede linear geordnete Teilmenge von  $Z$  hat eine obere Schranke, nämlich die Menge, aufgefaßt als Familie, aller Maße, die zu einer beliebigen in der linear geordneten Teilmenge enthaltenen Familie gehören. Nach dem Zornschen Lemma hat  $Z$  mindestens ein maximales Element.

$(P_{j_i}^{S_i})_{i \in I}$  sei ein maximales Element aus  $Z$ . Wir behaupten:  
 $(P_{j_i}^{S_i})_{i \in I}$  ist äquivalent zu  $(P_j)_{j \in \Theta}$ .

Zum Beweis genügt es, die Annahme, es gäbe ein  $A \in \mathcal{A}$  und ein  $j_0 \in \Theta$  mit  $P_{j_0}(A) > 0$  und  $P_{j_i}^{S_i}(A) = 0$  für alle  $i \in I$ , zu einem Widerspruch zu führen. Dieser ergibt sich wie folgt:

$$0 = P_{j_i}^{S_i}(A) = \int 1_A dP_{j_i}^{S_i} = \int_{S_i} 1_A dP_{j_i} = \int_{S_i} E(1_A | \mathcal{F}) dP_{j_i}$$

für alle  $i \in I$  zieht nach sich

$$P_{j_i}^{S_i}(\{E(1_A | \mathcal{F}) > 0\}) = 0 \text{ für alle } i \in I, \text{ und}$$

$$0 < P_{j_0}(A) = \int 1_A dP_{j_0} = \int E(1_A | \mathcal{F}) dP_{j_0} \text{ liefert}$$

$$P_{j_0}(\{E(1_A | \mathcal{F}) > 0\}) > 0.$$

Nun ist aber  $A' := \{E(1_A | \mathcal{F}) > 0\} \in \mathcal{F}$ , und man kann  $(P_{j_i}^{S_i})_{i \in I}$  mit  $P_{j_0}^{A'}$  echt vergrößern, was ein Widerspruch zur Maximalität von  $(P_{j_i}^{S_i})_{i \in I}$  ist.

Wir setzen  $P_i^{\mathcal{F}} := \Pr_{P_{j_i}^{S_i}}(P_{j_i})$  und behaupten: Für jedes  $j_0 \in \Theta$  sind höchstens abzählbar viele der Maße  $P_i^{\mathcal{F}}$ ,  $i \in I$ , vom Nullmaß verschieden.

Um das zu zeigen, nehmen wir an, die Aussage sei falsch; dann gibt es ein  $\varepsilon > 0$  mit  $P_{i_m}^{\mathcal{F}}(X) \geq \varepsilon$  für abzählbar unendlich viele  $i_m \in I$ ,  $m = 1, 2, \dots$  und  $i_m \neq i_{m'}$  für  $m \neq m'$ .

Nach (1.8) ist die Folge  $(P_{i_m}^{\mathcal{F}})_{m \in \mathbb{N}}$  paarweise orthogonal.

Wie in (1.5) sei  $(A_n)_{n \in \mathbb{N}}$  eine Zerlegung von  $X$  mit

$$P_{i_m}^{\mathcal{F}}(A_n) = 0 \text{ für } m \neq n, \text{ dann ist}$$

$$1 = P_{j_0}(X) = \sum_{n=1}^{\infty} P_{j_0}(A_n) \geq \sum_{n=1}^{\infty} P_{i_n}^{j_0}(A_n) = \sum_{n=1}^{\infty} P_{i_n}^{j_0}(X),$$

und wir haben einen Widerspruch, da die letzte Summe nicht endlich ist.

$J, \emptyset \neq J \subset I$ , sei also eine abzählbare Menge mit  $P_i^{j_0} = 0$  für alle  $i \in I \setminus J$ . Wir behaupten nun:

$P_{j_0}$  ist bezüglich  $(P_i^{j_0})_{i \in J}$  absolut stetig.

Wir zeigen, daß die Annahme, es gäbe ein  $A \in \mathcal{A}$  mit  $P_{j_0}(A) > 0$  und  $P_i^{j_0}(A) = 0$  für alle  $i \in I$ , zu einem Widerspruch führt.

( $P_i^{j_0}(A) = 0$  gilt sowieso für alle  $i \in I \setminus J$ , falls  $I \neq J$  ist.)

Wenn  $P_{j_0}(A) > 0$  ist, gibt es ein  $i_0 \in I$  mit  $P_{j_0}^{S_{i_0}}(A) > 0$ .

Aus der Definition von  $Pr_{P_{j_0}^{S_{i_0}}}(P_{j_0}) = P_{i_0}^{j_0}$  folgt dann

$P_{i_0}^{j_0}(A) > 0$ . Insbesondere ist  $i_0 \in J$ , und somit ein Wider-

spruch hergeleitet. Aus der Absolutstetigkeit von  $P_{j_0}$  bezüglich  $(P_i^{j_0})_{i \in J}$  folgt sofort die Absolutstetigkeit bezüglich  $(P_{j_0}^{S_{i_0}})_{i \in J}$ , was den Beweis von (1.9) vervollständigt.

Für das  $\sigma$  aus (1.9) kann man stets  $\mathcal{A}$  selbst nehmen, was im folgenden auch getan wird. Der nächste Satz beschreibt die Einbettung eines Experimentes in ein schwach dominiertes Experiment.

(1.10) Satz. Zu jedem Experiment  $(X, \mathcal{A}, (P_j)_{j \in \Theta})$  gibt es ein schwach dominiertes Experiment  $(X', \mathcal{A}', (P'_j)_{j \in \Theta})$  mit einem zu  $(P'_j)_{j \in \Theta}$  äquivalenten schwach dominierenden Maß  $\mu$  und eine  $\mathcal{A}'$ - $\mathcal{A}$ -meßbare Abbildung  $pr: X' \rightarrow X$  mit den Eigenschaften

(1)  $P_j = P'_j \circ pr^{-1}$  für alle  $j \in \Theta$ .

(2) Zu jedem  $f' \in \mathcal{L}^\infty(\mathcal{A}', \mu)$  gibt es ein gerichtetes System  $(f_\gamma)_{\gamma \in \Gamma}$  von beschränkten  $\mathcal{A}$ -meßbaren Funktionen so, daß es

für jedes  $g' \in \mathcal{X}^1(\mathcal{A}', \mu)$  ein  $\gamma_0 \in \Gamma$  gibt mit

$$\int |f' - f_{\gamma} \circ \text{pr}| g' d\mu = 0 \quad \text{für alle } \gamma \geq \gamma_0.$$

(3) Zu jedem  $A' \in \mathcal{A}'$  gibt es ein gerichtetes System  $(A_{\gamma})_{\gamma \in \Gamma}$  von Mengen  $A_{\gamma} \in \mathcal{A}$ ,  $\gamma \in \Gamma$ , so, daß es für jedes  $j \in \Theta$  ein  $\gamma_0 \in \Gamma$  gibt mit

$$P'_j(\text{pr}^{-1}(A_{\gamma}) \Delta A') = 0 \quad \text{für alle } \gamma \geq \gamma_0.$$

Beweis.  $(P_{j_i}^{S_i})_{i \in I}$  sei die Familie aus (1.9), wobei  $\mathcal{S} = \mathcal{A}$  gesetzt werde. Für alle  $i \in I$  sei

$$X_i := \{(x, i) : x \in X\} \text{ und } \mathcal{A}_i := \{A \times \{i\} : A \in \mathcal{A}\}.$$

$(X_i, \mathcal{A}_i)$  ist für alle  $i \in I$  ein Meßraum. Wir definieren

$$X' := \bigcup_{i \in I} X_i \quad \text{und } \mathcal{A}' \text{ als die direkte Summe der } \mathcal{A}_i, i \in I.$$

$(X', \mathcal{A}')$  ist ein Meßraum, und die Elemente  $A' \in \mathcal{A}'$  haben die Gestalt  $A' = \bigcup_{i \in I} A_i$ ,  $A_i \in \mathcal{A}_i$  für alle  $i \in I$ .

Für alle  $i \in I$  sei das Maß  $P_{j_i}^{S_i}$  auf  $(X', \mathcal{A}')$  definiert durch

$$P_{j_i}^{S_i}(A') := P_{j_i}^{S_i}(A), \text{ wenn } A' \cap X_i = A \times \{i\} \text{ ist, für alle } A' \in \mathcal{A}'.$$

Das Maß  $\mu$  auf  $(X', \mathcal{A}')$  sei definiert durch

$$\mu(A') := \begin{cases} \sum_{j \in J} P_{j_j}^{S_j}(A'), & J := \{i \in I : P_{j_i}^{S_i}(A') > 0, \text{ wenn } J \\ & \text{abzählbar und nicht leer ist und} \\ & \text{die Summe konvergiert.} \\ 0, & \text{wenn } J \text{ leer ist.} \\ \infty & \text{sonst.} \end{cases}$$

$\mu$  ist auf  $\mathcal{A}'$  lokalisierbar und hat die endliche Teilmengeneigenschaft. Das ergibt sich aus [15], S. 263 Theorem 1.

Die Abbildung  $\text{pr}$  sei folgendermaßen definiert:

$$\text{pr}: X' \ni (x, i) \mapsto x \in X.$$

$pr$  ist  $\alpha'$ - $\alpha$ -meßbar, und es ist  $P_{j_1}^{S_i} = P_{j_1}^{S_i} \circ pr^{-1}$  für alle  $i \in I$ .

Aus (1.9) folgt, daß es zu einem festen  $j_0 \in J$  ein Maß der Gestalt  $\sum_{j \in J} P_{j_0}^{S_j}$ ,  $\emptyset \neq J \subset I$  und  $J$  abzählbar, gibt, bezüglich dessen  $P_{j_0}$  absolut stetig ist. Aus (1.5) folgt, daß  $\sum_{j \in J} P_{j_0}^{S_j}$   $\sigma$ -endlich ist.

Also hat  $P_{j_0}$  eine Dichte  $p_{j_0} \geq 0$  bezüglich  $\sum_{j \in J} P_{j_0}^{S_j}$ .

Auf  $\alpha'$  definieren wir nun das normierte Maß  $P'_{j_0}$  vermöge:

$$P'_{j_0}(A') := \int_A p_{j_0} \circ pr^{-1} \cdot \bigcup_{j \in J} X_j \, d\mu = \sum_{j \in J} \int_A p_{j_0} \circ pr \, dP'_{j_0}^{S_j}$$

für alle  $A' \in \alpha'$ .

Für alle  $A \in \alpha$  gilt

$$P'_{j_0}(pr^{-1}(A)) = \sum_{j \in J} \int_{pr^{-1}(A)} p_{j_0} \circ pr \, dP'_{j_0}^{S_j} = \sum_{j \in J} \int_A p_{j_0} \, dP_{j_0}^{S_j} = P_{j_0}(A).$$

Damit folgt (1).

$f'$  sei eine  $\alpha'$ -meßbare Funktion. Es existiere ein  $E \in \alpha'$  mit  $\sigma$ -endlichem  $\mu$ -Maß so, daß  $\mu((X' \setminus E) \cap \{f' \neq 0\}) = 0$  ist. Dann gibt es eine abzählbare Teilmenge  $J \subset I$ ,  $J \neq \emptyset$ , mit  $E \subset \bigcup_{j \in J} X_j$   $\mu$ -fast. Also gilt auch  $\mu((X' \setminus \bigcup_{j \in J} X_j) \cap \{f' \neq 0\}) = 0$ .  $(A_j)_{j \in J}$  sei die gemäß (1.5) existierende Zerlegung von  $X$  mit  $P_{j_0}^{S_j}(A_{j'}) = 0$  für  $j \neq j'$ .

Zu  $f'$  definieren wir eine  $\alpha$ -meßbare Funktion  $f$  durch

$$(*) \quad f(x) := \sum_{j \in J} f'(x, j) \cdot 1_{A_j \times \{j\}}(x, j) \quad \text{für alle } x \in X.$$

Für ein festes  $x_0 \in X$  ist immer nur ein Summand der Summe von Null verschieden. Beachtet man die Definition von  $P_{j_0}^{S_j}$  und  $A_j$ ,  $j \in J$ , so ergibt sich

$$f \circ pr = f' \left( \sum_{j \in J} P_{j_0}^{S_j} \right)\text{-fast überall.}$$

Ist das  $f'$   $\mu$ -integrierbar (das  $E$  existiert dann von selbst),

so gilt

$$\begin{aligned} \int f' d\mu &= \sum_{j \in J} \int f'_+ dP'_{\mathcal{A}_j^{S_j}} - \sum_{j \in J} \int f'_- dP'_{\mathcal{A}_j^{S_j}} = \sum_{j \in J} \int f' dP'_{\mathcal{A}_j^{S_j}} \\ &= \sum_{j \in J} \int f \circ pr dP'_{\mathcal{A}_j^{S_j}} = \sum_{j \in J} \int f dP'_{\mathcal{A}_j^{S_j}}. \end{aligned}$$

Seien  $g'_\nu$ ,  $\nu = 1, \dots, m$ ,  $\mu$ -integrierbare Funktionen, dann hat

$$E := \{x' \in X' : g'_\nu(x') \neq 0 \text{ für ein } \nu \in \{1, \dots, m\}\}$$

$\sigma$ -endliches  $\mu$ -Maß.

$h'$  sei aus  $\mathcal{L}^\infty(\mathcal{A}', \mu)$ .  $h' \cdot 1_E$  hat die Eigenschaften des obigen  $f'$ . Gemäß (\*) können wir  $h' \cdot 1_E$  eine  $\mathcal{A}$ -meßbare Funktion  $h$  zuordnen.  $J \subset I$  sei wieder die abzählbare Menge mit

$E \subset \bigcup_{j \in J} X_j$   $\mu$ -fast. Es gilt

$$\int |h' - h \circ pr| g'_\nu d\mu = \sum_{j \in J} \int |h' - h \circ pr| g'_\nu dP'_{\mathcal{A}_j^{S_j}} = 0$$

für  $\nu = 1, \dots, m$ .

$h$  kann dabei noch o.B.d.A. beschränkt angenommen werden.

Wir definieren

$$\Gamma := \{ \{g'_1, \dots, g'_m\} : g'_\nu \mu\text{-integrierbar, } \nu = 1, \dots, m, m \in \mathbb{N} \}.$$

Auf  $\Gamma$  wird eine teilweise Ordnung  $\leq$  definiert durch

$$\{g'_1, \dots, g'_m\} \leq \{g'_1, \dots, g'_m\}, \text{ wenn } \{g'_1, \dots, g'_m\} \subset \{g'_1, \dots, g'_m\}.$$

Nach dem Vorangehenden gibt es nun für jedes  $f' \in \mathcal{L}^\infty(\mathcal{A}', \mu)$

für jede Wahl von  $\{g'_1, \dots, g'_m\} \in \Gamma$  ein beschränktes  $\mathcal{A}$ -meßbares  $f_{\{g'_1, \dots, g'_m\}}$  mit

$$\int |f' - f_{\{g'_1, \dots, g'_m\}} \circ pr| g'_\nu d\mu = 0 \text{ für } \nu = 1, \dots, m.$$

$\Gamma$  ist gerichtet. Man kann daher das eben Gesagte folgender-

maßen ausdrücken: Für jedes  $f' \in \mathcal{L}^\infty(\mathcal{A}', \mu)$  gibt es ein gerichtetes System von beschränkten  $\mathcal{A}$ -meßbaren Funktionen

$(f_\gamma)_{\gamma \in \Gamma}$  so, daß für jedes  $g' \in \mathcal{L}^1(\mathcal{A}', \mu)$  ein  $\gamma_0 \in \Gamma$  existiert mit

$$\int |f' - f_\gamma \circ pr| g' d\mu = 0 \text{ für alle } \gamma \geq \gamma_0.$$

Damit hat man (2).

(3) ist ein Spezialfall von (2).

(1.10)(2) beschreibt, in welchem Sinne die Menge der Funktionen der Gestalt  $f \circ p_r$ , wobei  $f$   $\mathcal{A}$ -meßbar und beschränkt ist, in  $\mathcal{L}^\infty(\mathcal{A}', \mu)$  dicht liegt. Das Ergebnis entspricht dem von Le Cam [11], der das Problem funktionalanalytisch behandelt.

§ 2. Eigenschaften suffizienter Unter- $\sigma$ - Algebren für schwach dominierte Familien von normierten Maßen

$(X, \mathcal{A}, (P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}})$  sei ein schwach dominiertes Experiment. Das schwach dominierende Maß wird auf einer Unter- $\sigma$ -Algebra von  $\mathcal{A}$  im allgemeinen die Lokalisierbarkeit und die endliche Teilmengeneigenschaft verlieren. Es ist das Ziel dieses Paragraphen zu zeigen, daß im Falle einer suffizienten Unter- $\sigma$ -Algebra  $\mathcal{O}$  von  $\mathcal{A}$  das schwach dominierende Maß so wählbar ist, daß es die obigen Eigenschaften auf  $\mathcal{O}$  behält.

(2.1) Lemma.  $(X, \mathcal{A}, (P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}})$  sei ein schwach dominiertes Experiment. Dann ist das schwach dominierende Maß zu  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$  äquivalent wählbar.

Beweis. Das schwach dominierende Maß sei  $\mu$ . Wir definieren

$$T_{\mathfrak{A}} := \{x \in X : \frac{dP_{\mathfrak{A}}}{d\mu}(x) > 0\}. \text{ Es existiert } \sup_{\mathfrak{A} \in \mathcal{O}} \tilde{T}_{\mathfrak{A}}^{\mu}.$$

Sei  $T \in \sup_{\mathfrak{A} \in \mathcal{O}} \tilde{T}_{\mathfrak{A}}^{\mu}$  gewählt. Dann ist auch  $\mu^T$  lokalisierbar und hat die endliche Teilmengeneigenschaft. Die Existenz von Dichten der  $P_{\mathfrak{A}}$  bezüglich  $\mu^T$  ist klar; man kann nämlich bis auf jeweils eine  $\mu$ -Nullmenge die Dichten der  $P_{\mathfrak{A}}$  bezüglich  $\mu$  nehmen.

Zu zeigen bleibt, daß  $\mu^T$  bezüglich  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$  absolut stetig ist. Dies gilt, wenn für jedes  $A \subset T$ ,  $A \in \mathcal{A}$ , mit  $P_{\mathfrak{A}}(A) = 0$  für alle  $\mathfrak{A} \in \mathcal{O}$  gilt  $\mu^T(A) = 0$ .

Nun folgt aus  $P_{\mathfrak{A}}(A) = 0$  für alle  $\mathfrak{A} \in \mathcal{O}$ , daß  $\mu(A \cap T_{\mathfrak{A}}) = 0$  ist für alle  $\mathfrak{A} \in \mathcal{O}$ . Dann gilt  $T_{\mathfrak{A}} \subset T \setminus A$   $\mu$ -fast.

Wegen der Minimalität von  $T$  folgt  $\mu^T(A) = \mu(A) = 0$ .

Wir können jetzt also immer o.E.d.A. annehmen, daß das schwach dominierende Maß zu  $(P_{\mathfrak{A}})_{\mathfrak{A} \in \mathcal{O}}$  äquivalent ist.

(2.2) Lemma.  $(X, \mathcal{A}, (P_{\mathcal{J}})_{\mathcal{J} \in \Theta})$  sei ein Experiment, das durch das zu  $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$  äquivalente Maß  $\mu$  schwach dominiert sei.

$\mathcal{J}$  sei eine für  $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$  suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

Dann sind die Dichten  $\frac{dP_{\mathcal{J}}}{d\mu}$  so wählbar, daß

$T_{\mathcal{J}} := \{x \in X : \frac{dP_{\mathcal{J}}}{d\mu}(x) > 0\} \in \mathcal{J}$  ist für alle  $\mathcal{J} \in \Theta$ .

Beweis. Für jedes  $A \in \mathcal{A}$  sei  $E(1_A | \mathcal{J})$  ein bedingter  $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{J}$ . Sei  $\mathcal{J}_0 \in \Theta$ , dann ist

$$\int E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) \cdot 1_{T_{\mathcal{J}_0}} dP_{\mathcal{J}_0} = \int E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) dP_{\mathcal{J}_0} = \int 1_{T_{\mathcal{J}_0}} dP_{\mathcal{J}_0} = 1.$$

Weil  $0 \leq E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) \leq 1$   $\mu$ -fast überall ist, ergibt sich

$E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) \cdot 1_{T_{\mathcal{J}_0}} \leq 1_{T_{\mathcal{J}_0}}$   $\mu$ -fast überall.

Mit der obigen Gleichung folgt

$$0 = \int 1_{T_{\mathcal{J}_0}} \cdot (1 - E(1_{T_{\mathcal{J}_0}} | \mathcal{J})) \cdot \frac{dP_{\mathcal{J}_0}}{d\mu} d\mu.$$

Da der Integrand  $\mu$ -fast überall nicht negativ ist, hat man

$1_{T_{\mathcal{J}_0}} \cdot (1 - E(1_{T_{\mathcal{J}_0}} | \mathcal{J})) \cdot \frac{dP_{\mathcal{J}_0}}{d\mu} = 0$   $\mu$ -fast überall, also ist

$1_{T_{\mathcal{J}_0}} = E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) \cdot 1_{T_{\mathcal{J}_0}}$   $\mu$ -fast überall.

Insbesondere ist  $E(1_{T_{\mathcal{J}_0}} | \mathcal{J})$  auf  $T_{\mathcal{J}_0}$ , abgesehen von einer  $\mu$ -Nullmenge, gleich 1, wegen  $0 \leq E(1_{T_{\mathcal{J}_0}} | \mathcal{J})$   $\mu$ -fast überall gilt daher  $1_{T_{\mathcal{J}_0}} \leq E(1_{T_{\mathcal{J}_0}} | \mathcal{J})$   $\mu$ -fast überall.

Weiter hat man für alle  $\mathcal{J} \in \Theta$

$\int (1_{T_{\mathcal{J}_0}} - E(1_{T_{\mathcal{J}_0}} | \mathcal{J})) dP_{\mathcal{J}} = 0$ , also  $1_{T_{\mathcal{J}_0}} = E(1_{T_{\mathcal{J}_0}} | \mathcal{J})$   $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$ -fast überall bzw.  $1_{T_{\mathcal{J}_0}} = E(1_{T_{\mathcal{J}_0}} | \mathcal{J})$   $\mu$ -fast überall.

Daher kann man jedes  $\frac{dP_{\mathcal{J}_0}}{d\mu}$  auf einer  $\mu$ -Nullmenge so abändern, daß diese neue Dichte genau für die  $x \in \{E(1_{T_{\mathcal{J}_0}} | \mathcal{J}) > 0\}$  positiv ist.

(2.3) Lemma.  $(X, \mathcal{A}, (P_{\mathcal{J}})_{\mathcal{J} \in \Theta})$  sei ein Experiment, das durch das zu  $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$  äquivalente Maß  $\mu$  schwach dominiert sei.

$\mathcal{J}$  sei eine für  $(P_{\mathcal{J}})_{\mathcal{J} \in \Theta}$  suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

Die Dichten  $\frac{dP_\vartheta}{d\mu}$  seien so gewählt, daß  $T_\vartheta := \{x \in X: \frac{dP_\vartheta}{d\mu}(x) > 0\}$  in  $\mathcal{F}$  ist für alle  $\vartheta \in \Theta$ .

Dann gibt es eine Familie  $(X_i)_{i \in I}$ ,  $I \neq \emptyset$ ,  $\mathcal{F}$ -meßbarer Mengen derart, daß jedes  $X_{i_0}$ ,  $i_0 \in I$ , in einem  $T_{\vartheta_0}$ ,  $\vartheta_0 \in \Theta$ , enthalten ist und positives ( $\sigma$ -endliches)  $\mu$ -Maß hat,  $\mu(X_i \cap X_{i'}) = 0$  ist für alle  $i, i' \in I$  mit  $i \neq i'$  und aus  $\mu(A \cap X_i) = 0$ ,  $A \in \mathcal{A}$ , für alle  $i \in I$  folgt  $\mu(A) = 0$ .

Beweis.  $\mathcal{F}$  sei die Menge aller Familien  $\mathcal{F}$ -meßbarer Mengen  $(X_j)_{j \in J}$ ,  $J \neq \emptyset$ , mit  $\mu(X_j) > 0$  für alle  $j \in J$  und  $\mu(X_j \cap X_{j'}) = 0$  für alle  $j, j' \in J$  mit  $j \neq j'$  und mit der Eigenschaft, daß zu jedem  $j_0 \in J$  ein  $\vartheta_0 \in \Theta$  existiert mit  $X_{j_0} \subset T_{\vartheta_0}$ . Familien, die aufgefaßt als Mengen übereinstimmen, seien dabei nicht unterschieden.

$\mathcal{F}$  ist nicht leer, denn für ein  $\vartheta_0 \in \Theta$  ist  $(T_{\vartheta_0}) \in \mathcal{F}$ .

Auf  $\mathcal{F}$  wird eine teilweise Ordnung  $\leq$  definiert durch

$(X_j)_{j \in J} \leq (X_{j'})_{j' \in J'}$ , wenn für jedes  $j_0 \in J$  ein  $j'_0 \in J'$  existiert mit  $X_{j_0} = X_{j'_0}$ .

Jede linear geordnete Teilmenge von  $\mathcal{F}$  hat eine obere Schranke in  $\mathcal{F}$ , nämlich die Menge, aufgefaßt als Familie, die aus allen  $\mathcal{F}$ -meßbaren Mengen besteht, die in irgendeiner Familie der linear geordneten Teilmenge vorkommen. Nach dem Zornschen Lemma gibt es also eine maximale Familie in  $\mathcal{F}$ , die mit  $(X_i)_{i \in I}$  bezeichnet werde.

Die Annahme, es gäbe ein  $A \in \mathcal{A}$  mit  $\mu(A) > 0$ , aber  $\mu(A \cap X_i) = 0$  für alle  $i \in I$ , führen wir zum Widerspruch.

Wenn  $\mu(A) > 0$  ist, gibt es ein  $\vartheta_0 \in \Theta$  mit  $P_{\vartheta_0}(A) > 0$ .

Man hat dann

$0 < P_{\mathcal{J}_0}(A) = \int 1_A dP_{\mathcal{J}_0} = \int E(1_A | \mathcal{J}) dP_{\mathcal{J}_0} = \int E(1_A | \mathcal{J}) \cdot 1_{T_{\mathcal{J}_0}} dP_{\mathcal{J}_0}$ ,  
und daraus folgt  $\mu(\{E(1_A | \mathcal{J}) \cdot 1_{T_{\mathcal{J}_0}} > 0\}) > 0$ .

Andererseits ist wegen  $\mu(A \cap X_i) = 0$  und  $1_{X_i} \cdot 1_A \geq 0$  für alle  $i \in I$

$$\begin{aligned} 0 &= P_{\mathcal{J}}(X_i \cap A) = \int 1_{X_i} \cdot 1_A dP_{\mathcal{J}} \geq \int 1_{X_i} \cdot 1_A \cdot 1_{T_{\mathcal{J}_0}} dP_{\mathcal{J}} \\ &= \int E(1_{X_i} \cdot 1_A \cdot 1_{T_{\mathcal{J}_0}} | \mathcal{J}) dP_{\mathcal{J}} = \int 1_{X_i} \cdot 1_{T_{\mathcal{J}_0}} \cdot E(1_A | \mathcal{J}) dP_{\mathcal{J}} \geq 0 \end{aligned}$$

für alle  $i \in I$  und  $\mathcal{J} \in \mathcal{O}$ .

Daraus folgt  $1_{X_i} \cdot 1_{T_{\mathcal{J}_0}} \cdot E(1_A | \mathcal{J}) = 0$   $\mu$ -fast überall für alle  $i \in I$ .  
Dann ist  $\mu(X_i \cap \{1_{T_{\mathcal{J}_0}} \cdot E(1_A | \mathcal{J}) > 0\}) = 0$  für alle  $i \in I$ . Da  $\{1_{T_{\mathcal{J}_0}} \cdot E(1_A | \mathcal{J}) > 0\} \in \mathcal{J}$  ist, hat man nun einen Widerspruch zur Maximalität von  $(X_i)_{i \in I}$ .

Jedes  $X_i$ ,  $i \in I$ , hat  $\sigma$ -endliches  $\mu$ -Maß, da jedes  $T_{\mathcal{J}}$ ,  $\mathcal{J} \in \mathcal{O}$ ,  $\sigma$ -endliches  $\mu$ -Maß hat.

Wir haben jetzt die Mittel bereitgestellt, um ein spezielles schwach dominierendes Maß konstruieren zu können.

(2.4) Satz.  $(X, \mathcal{A}, (P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}})$  sei ein Experiment, das durch das zu  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$  äquivalente Maß  $\mu$  schwach dominiert sei.

$\mathcal{J}$  sei eine für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$  suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

Die Dichten  $\frac{dP_{\mathcal{J}}}{d\mu}$  seien so gewählt, daß für alle  $\mathcal{J} \in \mathcal{O}$

$$T_{\mathcal{J}} := \{x \in X : \frac{dP_{\mathcal{J}}}{d\mu}(x) > 0\} \in \mathcal{J} \text{ ist.}$$

Dann gibt es eine Familie  $(X_i)_{i \in I}$ ,  $I \neq \emptyset$ ,  $\mathcal{J}$ -meßbarer Mengen und zu jedem  $i \in I$  ein  $\mathcal{J}_i \in \mathcal{O}$  so, daß jedes  $X_i$ ,  $i \in I$ , positives  $\sigma$ -endliches  $\mu$ -Maß hat und in  $T_{\mathcal{J}_i}$  liegt,

$\mu(X_i \cap X_{i'}) = 0$  ist für alle  $i, i' \in I$  mit  $i \neq i'$  und aus  $\mu(A \cap X_i) = 0$ ,  $A \in \mathcal{A}$ , für alle  $i \in I$  folgt  $\mu(A) = 0$ .

Weiter gibt es ein zu  $\mu$  äquivalentes schwach dominierendes Maß  $\nu$  so, daß  $\nu|_{\mathcal{J}}$  lokalisierbar ist und die endliche Teil-

mengeneigenschaft hat und für jedes  $i \in I$  gilt  $\nu^{X_i} = P_{\mathcal{F}_i}^{X_i}$ .

Beweis. Die Familie  $(X_i)_{i \in I}$  gibt es nach (2.3). Ferner ist  $\mu^{X_i}$  äquivalent zu  $P_{\mathcal{F}_i}^{X_i}$ .

Das Maß  $\nu$  auf  $(X, \mathcal{A})$  wird nun wie folgt definiert:

Für  $A \subset X_i$ ,  $A \in \mathcal{A}$ , sei  $\nu(A) := P_{\mathcal{F}_i}^{X_i}(A)$ ,  $i \in I$ .

Für beliebiges  $A \in \mathcal{A}$  sei

$$\nu(A) := \begin{cases} \sum_{i \in J} \nu(A \cap X_i), & J := \{i \in I : P_{\mathcal{F}_i}^{X_i}(A \cap X_i) > 0\}, \text{ wenn} \\ & J \text{ abzählbar aber nicht leer ist und} \\ & \text{die Summe konvergiert.} \\ 0, & \text{wenn } J \text{ leer ist.} \\ \infty & \text{sonst.} \end{cases}$$

$\nu$  ist wohldefiniert, denn sei  $X_{i_0} \cap X_{j_0} \neq \emptyset$  für  $i_0 \neq j_0$ , dann gilt wegen  $\mu(X_{i_0} \cap X_{j_0}) = 0$

$\mu^{X_{i_0}}(X_{i_0} \cap X_{j_0}) = 0$  und  $\mu^{X_{j_0}}(X_{i_0} \cap X_{j_0}) = 0$  und damit auch  $P_{\mathcal{F}_{i_0}}^{X_{i_0}}(X_{i_0} \cap X_{j_0}) = 0$  und  $P_{\mathcal{F}_{j_0}}^{X_{j_0}}(X_{i_0} \cap X_{j_0}) = 0$ .

Definitionsgemäß hat  $\nu$  die endliche Teilmengeneigenschaft.

Da  $\nu^{X_i} = P_{\mathcal{F}_i}^{X_i}$  ein endliches Maß ist und für alle  $i \in I$   $X_i \in \mathcal{O}$  ist, hat auch  $\nu|_{\mathcal{O}}$  die endliche Teilmengeneigenschaft.

Die Mengen  $T_{\mathcal{F}}$ ,  $\mathcal{F} \in \mathcal{O}$ , haben  $\sigma$ -endliches  $\mu$ -Maß. Deswegen gilt für ein  $\mathcal{F}_0 \in \mathcal{O}$ , daß  $\mu(T_{\mathcal{F}_0} \cap X_i)$  nur für abzählbar viele  $i \in I$  von Null verschieden ist. Dies ergibt sich wie in [15], S. 259 Lemma  $\alpha$ . Dann folgt aber, daß die Mengen  $T_{\mathcal{F}}$ ,  $\mathcal{F} \in \mathcal{O}$ ,  $\sigma$ -endliches  $\nu$ -Maß haben nach der Definition von  $\nu$ . Nun ist  $P_{\mathcal{F}} = P_{\mathcal{F}}^{T_{\mathcal{F}}}$  bezüglich  $\nu^{T_{\mathcal{F}}}$ ,  $\mathcal{F} \in \mathcal{O}$ , absolut stetig. Nach dem Satz von Radon-Nikodym hat  $P_{\mathcal{F}}$  eine Dichte bezüglich  $\nu^{T_{\mathcal{F}}}$ ,  $\mathcal{F} \in \mathcal{O}$ . Daraus folgt, daß  $P_{\mathcal{F}}$ ,  $\mathcal{F} \in \mathcal{O}$ , auch eine Dichte bezüglich  $\nu$  hat.

Unmittelbar aus der Definition von  $\nu$  folgt die Äquivalenz von  $\mu$  und  $\nu$ . Daraus ergibt sich die Lokalisierbarkeit von  $\nu$  auf  $\mathcal{O}$ .

Es bleibt noch zu zeigen, daß  $\nu|\mathcal{O}$  lokalisierbar ist.

Sei  $(S_j)_{j \in J}$ ,  $J \neq \emptyset$ , eine Familie  $\mathcal{O}$ -meßbarer Mengen. Zunächst existiert  $\sup_{j \in J} \tilde{S}_j^\nu$ . Es genügt zu zeigen, daß ein  $S \in \mathcal{O}^*$  existiert mit  $S \in \sup_{j \in J} \tilde{S}_j^\nu$ .

Sei  $A \in \sup_{j \in J} \tilde{S}_j^\nu$ . Dann ist  $1_{S_j} \leq 1_A$   $\nu$ -fast überall für alle  $j \in J$ . Daraus folgt  $1_{S_j} = E(1_{S_j}|\mathcal{O}) \leq E(1_A|\mathcal{O})$   $\nu$ -fast überall für alle  $j \in J$ . Wegen der Minimalität von  $A$  ist dann  $1_A \leq E(1_A|\mathcal{O})$   $\nu$ -fast überall. Man hat also

$$\int (E(1_A|\mathcal{O}) - 1_A) dP_\vartheta = 0 \quad \text{für alle } \vartheta \in \mathcal{O},$$

wobei der Integrand  $\nu$ -fast überall nicht negativ ist.

Das liefert  $E(1_A|\mathcal{O}) = 1_A$   $(P_\vartheta)_{\vartheta \in \mathcal{O}}$ -fast überall, also  $E(1_A|\mathcal{O}) = 1_A$   $\nu$ -fast überall.

Die  $\mathcal{O}$ -meßbare Menge  $\{E(1_A|\mathcal{O}) > 0\}$  ist daher in  $\sup_{j \in J} \tilde{S}_j^\nu$  enthalten.

### § 3. Suffizienz und paarweise Suffizienz

(3.1) Definition. Sei  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  ein Experiment und  $\mathcal{G}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ . Dann heißt  $\mathcal{G}$  paarweise suffizient für  $(P_{\vartheta})_{\vartheta \in \Theta}$  bzw.  $(X, \mathcal{G}, (P_{\vartheta}|_{\mathcal{G}})_{\vartheta \in \Theta})$  paarweise suffizient, wenn für jede Wahl von  $\vartheta_1, \vartheta_2 \in \Theta$   $\mathcal{G}$  suffizient für  $\{P_{\vartheta_1}, P_{\vartheta_2}\}$  ist.

Ein Beispiel dafür, daß paarweise Suffizienz im allgemeinen nicht Suffizienz liefert, steht in [7].

Um beweisen zu können, daß die paarweise Suffizienz aber in gewissen Fällen die Suffizienz nach sich zieht, benötigen wir den Begriff des Querschnittes einer Familie von Äquivalenzklassen meßbarer Funktionen [15], S. 250.

(3.2) Definition.  $(X, \mathcal{A}, \mu)$  sei ein Maßraum.

(a) Eine Familie von Äquivalenzklassen  $\mathcal{A}$ -meßbarer Funktionen  $(\widetilde{f}_E)_{E \in \mathcal{A}, \mu(E) < \infty}$  heißt Querschnitt ("cross-section"), wenn für alle  $f \in \widetilde{f}_E$ , wobei  $E \in \mathcal{A}$  und  $\mu(E) < \infty$  ist,

$f \cdot 1_{X \setminus E} = 0$   $\mu$ -fast überall ist und für  $E, F \in \mathcal{A}$  mit  $\mu(E) < \infty$  und  $\mu(F) < \infty$  gilt  $\widetilde{f}_{E \cap F} = \widetilde{f_E \cdot 1_{E \cap F}} = \widetilde{f_F \cdot 1_{E \cap F}}$ .

(b) Ein Querschnitt heißt durch eine  $\mathcal{A}$ -meßbare Funktion  $f$  bestimmt, wenn  $\widetilde{f}_E = \widetilde{f \cdot 1_E}$  ist für alle  $E \in \mathcal{A}$  mit  $\mu(E) < \infty$ .

Ist  $\mu$   $\sigma$ -endlich, so ist jeder Querschnitt durch eine meßbare Funktion bestimmt. Für beliebige Maße trifft dies jedoch nicht immer zu.

(3.3) Definition.  $(X, \mathcal{A}, \mu)$  sei ein Maßraum und  $(\widetilde{f}_E)_{E \in \mathcal{A}, \mu(E) < \infty}$  ein Querschnitt. Dann sei für  $A \in \mathcal{A}$  mit  $A = \bigcup_{n=1}^{\infty} E_n$ , wobei  $(E_n)_{n \in \mathbb{N}}$  eine Folge paarweise disjunkter Mengen aus  $\mathcal{A}$  mit endlichem  $\mu$ -Maß ist,  $\widetilde{f}_A$  die Äquivalenzklasse der durch

$$f_A(x) := \begin{cases} \sum_{n=1}^{\infty} f_{E_n}(x) \cdot 1_{E_n}(x) & \text{für } x \in A, \text{ wobei } f_{E_n} \in \tilde{f}_{E_n} \text{ ist,} \\ 0 & \text{für } x \in X \setminus A \end{cases}$$

definierten  $\mathcal{A}$ -meßbaren Funktion.

Wir werden häufig folgenden Satz verwenden:

(3.4) Satz.  $(X, \mathcal{A}, \mu)$  sei ein Maßraum, und  $\mu$  habe die endliche Teilmengeneigenschaft.  $\mu$  ist genau dann lokalisierbar, wenn jeder Querschnitt durch eine  $\mathcal{A}$ -meßbare Funktion bestimmt ist.

Zum Beweis vergleiche man [15], S.264 Theorem 2.

(3.5) Satz.  $(X, \mathcal{A}, (P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}})$  sei ein Experiment und  $\mathcal{O}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

(a) Ist  $\mathcal{O}$  suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$ , so ist  $\mathcal{O}$  paarweise suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$ .

(b) Ist  $\mathcal{O}$  paarweise suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$  und ist  $(P_{\mathcal{J}}|\mathcal{O})_{\mathcal{J} \in \mathcal{O}}$  schwach dominiert durch  $\nu$  auf  $\mathcal{O}$ , so ist  $\mathcal{O}$  suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$ .

Beweis. (a) ist klar.

(b): Nach (2.1) können wir o.E.d.A. annehmen, daß  $\nu$  zu  $(P_{\mathcal{J}}|\mathcal{O})_{\mathcal{J} \in \mathcal{O}}$  äquivalent ist. Für jedes  $\mathcal{J} \in \mathcal{O}$  sei

$$T_{\mathcal{J}} := \{x \in X : \frac{dP_{\mathcal{J}}|\mathcal{O}}{d\nu}(x) > 0\}.$$

(\*) Für jedes  $\mathcal{J}_0 \in \mathcal{O}$  und  $S \in \mathcal{O}$  folgt dann aus  $P_{\mathcal{J}_0}(S \cap T_{\mathcal{J}_0}) = 0$ , daß  $\nu(S \cap T_{\mathcal{J}_0}) = 0$  und daher  $P_{\mathcal{J}}(S \cap T_{\mathcal{J}_0}) = 0$  für alle  $\mathcal{J} \in \mathcal{O}$  ist.

Für jedes  $S_0 \in \mathcal{O}$  mit  $\sigma$ -endlichem  $\nu$ -Maß ist  $((P_{\mathcal{J}}|\mathcal{O})^{S_0})_{\mathcal{J} \in \mathcal{O}}$  durch das  $\sigma$ -endliche Maß  $\nu^{S_0}$  dominiert, und es gibt eine abzählbare Teilmenge  $\mathcal{O}_{S_0} \subset \mathcal{O}$  so, daß  $((P_{\mathcal{J}}|\mathcal{O})^{S_0})_{\mathcal{J} \in \mathcal{O}}$  zu  $((P_{\mathcal{J}}|\mathcal{O})^{S_0})_{\mathcal{J} \in \mathcal{O}_{S_0}}$  äquivalent ist. Insbesondere gilt

$$S_0 \subset \bigcup_{\mathcal{J} \in \mathcal{O}_{S_0}} T_{\mathcal{J}} \quad \nu\text{-fast.}$$

Nach Voraussetzung existiert für jedes  $A \in \mathcal{A}$  und jede Wahl von  $\mathcal{V}_1, \mathcal{V}_2 \in \mathcal{O}$  ein bedingter  $\{P_{\mathcal{V}_1}, P_{\mathcal{V}_2}\}$ -Erwartungswert  $E_{\mathcal{V}_1, \mathcal{V}_2}(1_A | \mathcal{F})$  von  $1_A$  bezüglich  $\mathcal{F}$ . Es gilt also

$$\int_S 1_A dP_{\mathcal{V}_i} = \int_S E_{\mathcal{V}_1, \mathcal{V}_2}(1_A | \mathcal{F}) dP_{\mathcal{V}_i} | \mathcal{F} \quad \text{für alle } S \in \mathcal{F} \text{ und } i = 1, 2.$$

Daraus folgt

$$\int_S E_{\mathcal{V}, \mathcal{V}'}(1_A | \mathcal{F}) dP_{\mathcal{V}} | \mathcal{F} = \int_S E_{\mathcal{V}, \mathcal{V}''}(1_A | \mathcal{F}) dP_{\mathcal{V}} | \mathcal{F}$$

für alle  $A \in \mathcal{A}$ ,  $S \in \mathcal{F}$  und  $\mathcal{V}, \mathcal{V}', \mathcal{V}'' \in \mathcal{O}$ .

Es gilt daher, wenn man (\*) beachtet, für jedes  $A \in \mathcal{A}$  und  $\mathcal{V} \in \mathcal{O}$

$$E_{\mathcal{V}, \mathcal{V}'}(1_A | \mathcal{F}) \cdot 1_{T_{\mathcal{V}}} = E_{\mathcal{V}, \mathcal{V}''}(1_A | \mathcal{F}) \cdot 1_{T_{\mathcal{V}}} \quad \nu\text{-fast überall für alle } \mathcal{V}', \mathcal{V}'' \in \mathcal{O}.$$

Jetzt sei  $S_0$  aus  $\mathcal{F}$  mit positivem  $\sigma$ -endlichen  $\nu$ -Maß. Wir

wählen eine Abzählung  $(P_{\mathcal{V}_i} | \mathcal{F})^{S_0}$ ,  $i = 1, 2, \dots$ , von

$((P_{\mathcal{V}} | \mathcal{F})^{S_0})_{\mathcal{V} \in \mathcal{O}_{S_0}}$  ( $\mathcal{O}_{S_0}$  kann endlich, aber nicht leer sein;

es ist möglich, daß  $i$  nur den Wert 1 annimmt). Wir definie-

ren für jedes  $A \in \mathcal{A}$

$$\psi_1(1_A) := E_{\mathcal{V}_1, \mathcal{V}_1}(1_A | \mathcal{F}) \cdot 1_{T_{\mathcal{V}_1}} \quad \text{und für } i \geq 2$$

$$\psi_i(1_A) := \psi_{i-1}(1_A) \cdot 1_{\bigcup_{j=1}^{i-1} T_{\mathcal{V}_j}} + E_{\mathcal{V}_i, \mathcal{V}_i}(1_A | \mathcal{F}) \cdot 1_{T_{\mathcal{V}_i} \setminus \bigcup_{j=1}^{i-1} T_{\mathcal{V}_j}} \quad \text{und}$$

$$\psi_{S_0}(1_A) := 1_{S_0} \cdot \lim_{i \rightarrow \infty} \psi_i(1_A).$$

Wegen der oben gezeigten Eigenschaft von  $E_{\mathcal{V}, \mathcal{V}'}(1_A | \mathcal{F}) \cdot 1_{T_{\mathcal{V}}}$ ,

$A \in \mathcal{A}$  und  $\mathcal{V}, \mathcal{V}' \in \mathcal{O}$ , und da  $S_0 = (\bigcup_{\mathcal{V} \in \mathcal{O}_{S_0}} T_{\mathcal{V}}) \cap S_0$   $\nu$ -fast ist,

folgt

$$\int_{S \cap S_0} 1_A dP_{\mathcal{V}} = \int_{S \cap S_0} \psi_{S_0}(1_A) dP_{\mathcal{V}} | \mathcal{F} = \int_S \psi_{S_0}(1_A) dP_{\mathcal{V}} | \mathcal{F}$$

für alle  $A \in \mathcal{A}$ ,  $S \in \mathcal{F}$  und  $\mathcal{V} \in \mathcal{O}$ .

Aus dieser Gleichung ergibt sich, daß für jedes  $A \in \mathcal{A}$

$(\overline{\psi_S(1_A)})_{S \in \mathcal{F}, \nu(S) < \infty}$  ein Querschnitt von Äquivalenzklassen

$\mathcal{F}$ -meßbarer Funktionen ist, der wegen der Eigenschaften von  $\nu$

durch eine  $\mathcal{O}$ -meßbare Funktion  $\psi(1_A)$  bestimmt ist.

Für jedes  $S_0 \in \mathcal{O}$  mit  $\sigma$ -endlichem  $\nu$ -Maß gilt daher

$$\int_{S \cap S_0} 1_A dP_j = \int_{S \cap S_0} \psi(1_A) dP_j|_{\mathcal{O}} \quad \text{für alle } A \in \mathcal{A}, S \in \mathcal{O} \text{ und}$$

$j \in \mathcal{O}$ . Folglich ist für alle  $A \in \mathcal{A}$ ,  $S \in \mathcal{O}$  und  $j \in \mathcal{O}$

$$\int_S \psi(1_A) dP_j|_{\mathcal{O}} = \int_{S \cap T_j} \psi(1_A) dP_j|_{\mathcal{O}} = \int_{S \cap T_j} 1_A dP_j = \int_S 1_A dP_j,$$

woraus folgt, daß für jedes  $A \in \mathcal{A}$  die Funktion  $\psi(1_A)$  eine Version des gesuchten bedingten  $(P_j)_{j \in \mathcal{O}}$ -Erwartungswertes von  $1_A$  bezüglich  $\mathcal{O}$  ist.

Den Ansatz zur Konstruktion des  $\psi$  aus (3.5) entnehmen wir [14].

#### § 4. Ein verallgemeinertes Neyman-Kriterium

$(X, \mathcal{A}, (P_j)_{j \in \Theta})$  sei ein Experiment und  $\mathcal{D}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ . Wenn  $(P_j)_{j \in \Theta}$  (durch ein  $\sigma$ -endliches Maß auf  $\mathcal{A}$ ) dominiert ist, hat man das folgende Kriterium für die Suffizienz von  $\mathcal{D}$  für  $(P_j)_{j \in \Theta}$ .

(4.1) Satz. (Neyman-Kriterium) Hat jedes  $P_j$ ,  $j \in \Theta$ , eine Dichte bezüglich eines  $\sigma$ -endlichen Maßes  $\mu$  auf  $\mathcal{A}$ , so ist  $\mathcal{D}$  genau dann suffizient für  $(P_j)_{j \in \Theta}$ , wenn es eine  $\mathcal{A}$ -meßbare Funktion  $h$  und für jedes  $j \in \Theta$  eine  $\mathcal{D}$ -meßbare Funktion  $f_j$  gibt mit  $\frac{dP_j}{d\mu} = f_j \cdot h$   $\mu$ -fast überall.

Zum Beweis vgl. [1], S. 438 Korollar 6.1.

Wir wollen dies Kriterium für schwach dominierte Familien verallgemeinern. Dazu benötigen wir eine Erweiterung des Begriffes "bedingter Erwartungswert".

(4.2) Satz.  $(X, \mathcal{A}, \mu)$  sei ein Maßraum.  $\mathcal{D}$  sei eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ , und  $\mu|_{\mathcal{D}}$  sei lokalisierbar und habe die endliche Teilmengeneigenschaft.

Dann gibt es zu jeder  $\mathcal{A}$ -meßbaren Funktion  $f$ , für die  $\int_E f d\mu$  für alle  $E \in \mathcal{D}$  mit  $\mu(E) < \infty$  existiert und endlich ist (also insbesondere für  $f \in \mathcal{L}^1(\mathcal{A}, \mu)$  oder  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$ ), eine  $\mathcal{D}$ -meßbare Funktion  $E(f|\mathcal{D})$  mit

$$\int_E E(f|\mathcal{D}) d\mu = \int_E f d\mu \quad \text{für alle } E \in \mathcal{D} \text{ mit } \mu(E) < \infty.$$

$E(f|\mathcal{D})$  heißt bedingter  $\mu$ -Erwartungswert von  $f$  bezüglich  $\mathcal{D}$ .

Beweis. Zunächst existiert nach dem Satz von Radon-Nikodym ein Querschnitt von  $\mu|_{\mathcal{D}}$ -Äquivalenzklassen  $\mathcal{D}$ -meßbarer Funktionen  $(\tilde{g}_E^{\mu|_{\mathcal{D}}})_{E \in \mathcal{D}, \mu(E) < \infty}$  derart, daß für  $E \in \mathcal{D}$  mit  $\mu(E) < \infty$  gilt  $\int_{E'} \tilde{g}_E d\mu = \int_E f d\mu$  für alle  $E' \in \mathcal{D}$  mit  $E' \subset E$ .

Da  $\mu|_{\mathcal{O}}$  aber lokalisierbar ist und die endliche Teilmengeneigenschaft hat, ist der Querschnitt durch eine  $\mathcal{O}$ -meßbare Funktion  $g$  mit  $\widetilde{g \cdot 1_E}^{\mu|_{\mathcal{O}}} = \widetilde{g}_E^{\mu|_{\mathcal{O}}}$  für alle  $E \in \mathcal{O}$  mit  $\mu(E) < \infty$  bestimmt.  $E(f|\mathcal{O}) := g$  leistet daher das Gewünschte.

(4.3) Bemerkung. Unter den Voraussetzungen von (4.2) sei  $f$  eine  $\mathcal{A}$ -meßbare Funktion und  $g$  eine  $\mathcal{O}$ -meßbare Funktion so, daß  $E(g \cdot f|\mathcal{O})$  und  $E(f|\mathcal{O})$  existieren. Dann ist  $\mu$ -fast überall  $E(g \cdot f|\mathcal{O}) = g \cdot E(f|\mathcal{O})$ , wie man durch die Betrachtung von Querschnitten sieht. Es gilt also dieselbe Rechenregel wie für bedingte Erwartungswerte im herkömmlichen Sinn.

Bedingte Erwartungswerte im Zusammenhang mit nicht notwendig endlichen Maßen behandelt auch A. Ionescu Tulcea in [9], jedoch mit anderen Methoden.

(4.4) Satz.  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  sei ein Experiment und  $\mathcal{O}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

(a) Ist  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  schwach dominiert und  $\mathcal{O}$  suffizient für  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}}$ , so gibt es ein zu  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}}$  äquivalentes schwach dominierendes Maß  $\nu$  so, daß die Dichten  $\frac{dP_{\mathcal{V}}}{d\nu}$  für jedes  $\mathcal{V} \in \mathcal{O}$   $\mathcal{O}$ -meßbar wählbar sind und  $\nu|_{\mathcal{O}}$  für das Unterexperiment

$(X, \mathcal{O}, (P_{\mathcal{V}}|_{\mathcal{O}})_{\mathcal{V} \in \mathcal{O}})$  ein schwach dominierendes Maß ist.

(b) Existiert auf  $\mathcal{A}$  ein Maß  $\nu$  so, daß jedes  $P_{\mathcal{V}}$ ,  $\mathcal{V} \in \mathcal{O}$ , bezüglich  $\nu$  eine  $\mathcal{O}$ -meßbare Dichte hat und das Unterexperiment  $(X, \mathcal{O}, (P_{\mathcal{V}}|_{\mathcal{O}})_{\mathcal{V} \in \mathcal{O}})$  durch  $\nu|_{\mathcal{O}}$  schwach dominiert ist, so ist  $\mathcal{O}$  suffizient für  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}}$ .

Beweis. (a): Das schwach dominierende Maß kann nach (2.1) zu  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}}$  äquivalent gewählt werden. Es habe außerdem die in (2.4) angegebene spezielle Form. Wir benennen dies Maß wie in (2.4) mit  $\nu$  und verwenden auch sonst die dortigen

Bezeichnungen. Ist  $f$  eine  $\mathcal{A}$ -meßbare Funktion, die für jedes  $P_{\mathcal{J}}$ ,  $\mathcal{J} \in \mathcal{O}$ , integrierbar ist, so bezeichne  $E(f|\mathcal{J})$  einen bedingten  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{O}}$ -Erwartungswert von  $f$ . Zu  $E \in \mathcal{J}$  mit  $\nu(E) < \infty$  gibt es gemäß (2.4) eine abzählbare Teilmenge  $J \subset I$ ,  $J \neq \emptyset$ , mit  $E \subset \bigcup_{i \in J} X_i$   $\nu$ -fast. Es war  $X_i \in \mathcal{J}$  für alle  $i \in I$ . Daher gilt für eine nicht negative beschränkte  $\mathcal{A}$ -meßbare Funktion  $f$

$$\begin{aligned} \int_E f \, d\nu &= \int f \cdot 1_E \, d\nu = \sum_{i \in J} \int_E f \, d\nu^{X_i} = \sum_{i \in J} \int_E f \, dP_{\mathcal{J}_i}^{X_i} \\ &= \sum_{i \in J} \int_E E(f|\mathcal{J}) \, dP_{\mathcal{J}_i}^{X_i} = \int_E E(f|\mathcal{J}) \, d\nu \quad \text{für alle } E \in \mathcal{J} \text{ mit } \nu(E) < \infty. \end{aligned}$$

Also ist  $E(f|\mathcal{J})$  auch ein bedingter  $\nu$ -Erwartungswert von  $f$  bezüglich  $\mathcal{J}$ .

Es existieren  $\mathcal{J}$ -meßbare Dichten  $\frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}}$ ,  $\mathcal{J} \in \mathcal{O}$ , weil gemäß der Konstruktion von  $\nu$  die Mengen  $T_{\mathcal{J}} := \{x \in X : \frac{dP_{\mathcal{J}}}{d\nu}(x) > 0\}$   $\sigma$ -endliches  $\nu|\mathcal{J}$ -Maß haben für alle  $\mathcal{J} \in \mathcal{O}$ .

Für alle  $A \in \mathcal{A}$ ,  $E \in \mathcal{J}$  mit  $\nu(E) < \infty$  und  $\mathcal{J} \in \mathcal{O}$  gilt nach (4.3)

$$\int_E 1_A \cdot \frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}} \, d\nu = \int_E E(1_A|\mathcal{J}) \frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}} \, d\nu.$$

Dann gilt aber auch für alle  $A \in \mathcal{A}$  und  $\mathcal{J} \in \mathcal{O}$

$$\begin{aligned} P_{\mathcal{J}}(A) &= \int_{T_{\mathcal{J}}} 1_A \, dP_{\mathcal{J}} = \int_{T_{\mathcal{J}}} E(1_A|\mathcal{J}) \, dP_{\mathcal{J}} = \int_{T_{\mathcal{J}}} E(1_A|\mathcal{J}) \frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}} \, d\nu \\ &= \int_{T_{\mathcal{J}}} 1_A \cdot \frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}} \, d\nu, \end{aligned}$$

weil  $T_{\mathcal{J}}$  für alle  $\mathcal{J} \in \mathcal{O}$   $\sigma$ -endliches  $\nu|\mathcal{J}$ -Maß hat, und es ist

$$\frac{dP_{\mathcal{J}}}{d\nu} = \frac{dP_{\mathcal{J}}|\mathcal{J}}{d\nu|\mathcal{J}} \quad \nu\text{-fast überall.}$$

$\nu$  hat also die in (a) gewünschten Eigenschaften.

(b): Die Voraussetzungen liefern die Existenz eines bedingten  $\nu$ -Erwartungswertes  $E(f|\mathcal{J})$  für jede  $\mathcal{A}$ -meßbare Funktion  $f$ , für die  $\int_E f \, d\nu$  für alle  $E \in \mathcal{J}$  mit  $\nu(E) < \infty$  existiert und endlich ist.

$\frac{dP_{\mathcal{J}}}{d\nu}$  sei für alle  $\mathcal{J} \in \mathcal{O}$   $\mathcal{J}$ -meßbar gewählt. Es sei für alle

$\vartheta \in \Theta$   $T_{\vartheta} := \{x \in X : \frac{dP_{\vartheta}}{d\nu}(x) > 0\}$ . Diese Mengen haben  $\sigma$ -endliches  $\nu|_{\mathcal{X}}$ -Maß. Daher gilt

$$\begin{aligned} \int E(1_A | \mathcal{G}) dP_{\vartheta} &= \int_{T_{\vartheta}} E(1_A | \mathcal{G}) dP_{\vartheta} = \int_{T_{\vartheta}} E(1_A | \mathcal{G}) \frac{dP_{\vartheta}}{d\nu} d\nu \\ &= \int_{T_{\vartheta}} E(1_A \cdot \frac{dP_{\vartheta}}{d\nu} | \mathcal{G}) d\nu = \int_{T_{\vartheta}} 1_A \cdot \frac{dP_{\vartheta}}{d\nu} d\nu = \int 1_A dP_{\vartheta} \end{aligned}$$

für alle  $A \in \mathcal{A}$  und  $\vartheta \in \Theta$ . Daraus folgt die Suffizienz von  $\mathcal{G}$  für  $(P_{\vartheta})_{\vartheta \in \Theta}$ .

Die obige Beweisidee und auch die des ersten Teiles des folgenden verallgemeinerten Neyman-Kriteriums ist die gleiche wie bei der Dominiertheit durch ein  $\sigma$ -endliches Maß.

#### (4.5) Satz. (Verallgemeinertes Neyman-Kriterium)

$(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  sei ein Experiment und  $\mathcal{G}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$ .

(a) Ist  $(P_{\vartheta})_{\vartheta \in \Theta}$  schwach dominiert durch ein Maß  $\mu$  und  $\mathcal{G}$  suffizient für  $(P_{\vartheta})_{\vartheta \in \Theta}$ , so gibt es eine  $\mathcal{A}$ -meßbare Funktion  $h$  und für jedes  $\vartheta \in \Theta$  eine  $\mathcal{G}$ -meßbare Funktion  $f_{\vartheta}$  mit  $\frac{dP_{\vartheta}}{d\mu} = f_{\vartheta} \cdot h$  für alle  $\vartheta \in \Theta$   $\mu$ -fast überall. Weiter gibt es auf  $\mathcal{G}$  ein Maß, das  $(P_{\vartheta}|_{\mathcal{G}})_{\vartheta \in \Theta}$  schwach dominiert.

(b) Existiert auf  $\mathcal{A}$  ein Maß  $\mu$ , bezüglich dessen  $P_{\vartheta}$  für jedes  $\vartheta \in \Theta$  eine Dichte hat, und gibt es eine  $\mathcal{A}$ -meßbare Funktion  $h$  und für jedes  $\vartheta \in \Theta$  eine  $\mathcal{G}$ -meßbare Funktion  $f_{\vartheta}$  mit  $\frac{dP_{\vartheta}}{d\mu} = f_{\vartheta} \cdot h$  für alle  $\vartheta \in \Theta$   $\mu$ -fast überall und ist  $(P_{\vartheta}|_{\mathcal{G}})_{\vartheta \in \Theta}$  schwach dominiert durch  $\mu$  auf  $\mathcal{G}$ , so ist  $\mathcal{G}$  suffizient für  $(P_{\vartheta})_{\vartheta \in \Theta}$ .

Beweis. (a):  $\mu$  kann o.E.d.A. zu  $(P_{\vartheta})_{\vartheta \in \Theta}$  äquivalent angenommen werden. Wir haben dann die gleichen Voraussetzungen wie in (2.4).  $\nu$  sei das dort angegebene zu  $\mu$  äquivalente und für  $(P_{\vartheta})_{\vartheta \in \Theta}$  schwach dominierende Maß auf  $\mathcal{A}$ .  $\nu|_{\mathcal{X}}$  war schwach

dominierend für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ .

Sei  $A \in \mathcal{A}$  mit  $\nu(A) < \infty$  gegeben. Nach der Konstruktion von  $\nu$  hat  $A$  dann  $\sigma$ -endliches  $\mu$ -Maß. Nach dem Satz von Radon-Nikodym gibt es daher eine  $\mathcal{A}$ -meßbare Funktion  $h_A$ , die auf  $X \setminus A$  verschwindet, mit

$$\int_B h_A d\mu = \int_B d\nu \quad \text{für alle } B \in \mathcal{A} \text{ mit } B \subset A.$$

$(\tilde{h}_A^\nu)_{A \in \mathcal{A}, \nu(A) < \infty}$  ist dann ein Querschnitt wegen der Äquivalenz von  $\nu$  und  $\mu$ , der durch eine  $\mathcal{A}$ -meßbare Funktion  $h$  bestimmt ist. Es gilt

$$\int_A h d\mu = \int_A d\nu \quad \text{für alle } A \in \mathcal{A} \text{ mit } \nu(A) < \infty.$$

Daraus folgt, daß auch für jedes  $f \in \mathcal{L}^1(\mathcal{A}, \nu)$  gilt

$$\int f \cdot h d\mu = \int f d\nu.$$

Nach (4.4) hat jedes  $P_{\mathcal{J}}$ ,  $\mathcal{J} \in \mathcal{C}$ , eine  $\mathcal{J}$ -meßbare Dichte  $f_{\mathcal{J}}$  bezüglich  $\nu$ . Weil  $f_{\mathcal{J}} \in \mathcal{L}^1(\mathcal{A}, \nu)$  für alle  $\mathcal{J} \in \mathcal{C}$  ist, gilt

$$\int_A f_{\mathcal{J}} \cdot h d\mu = \int_A f_{\mathcal{J}} d\nu = \int_A dP_{\mathcal{J}} \quad \text{für alle } \mathcal{J} \in \mathcal{C} \text{ und } A \in \mathcal{A}.$$

Also ist  $\frac{dP_{\mathcal{J}}}{d\mu} = f_{\mathcal{J}} \cdot h$   $\mu$ -fast überall.

(b): Es sei  $T_{\mathcal{J}} := \{x \in X : \frac{dP_{\mathcal{J}}}{d\mu}(x) > 0\}$  für alle  $\mathcal{J} \in \mathcal{C}$ .

Für alle  $\mathcal{J}_1, \mathcal{J}_2 \in \mathcal{C}$  hat  $T_{\mathcal{J}_1} \cup T_{\mathcal{J}_2}$   $\sigma$ -endliches  $\mu$ -Maß.  $\mu^{T_{\mathcal{J}_1} \cup T_{\mathcal{J}_2}}$  ist also ein  $\sigma$ -endliches Maß auf  $\mathcal{A}$ , und  $f_{\mathcal{J}_i} \cdot h$  ist auch bezüglich  $\mu^{T_{\mathcal{J}_1} \cup T_{\mathcal{J}_2}}$  eine Dichte für  $P_{\mathcal{J}_i}$ ,  $i = 1, 2$ . Nach (4.1) ist  $\mathcal{J}$  dann suffizient für  $\{P_{\mathcal{J}_1}, P_{\mathcal{J}_2}\}$  mit beliebigen  $\mathcal{J}_1, \mathcal{J}_2 \in \mathcal{C}$ .  $\mathcal{J}$  ist also paarweise suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ . Nach (3.5)(b) ist  $\mathcal{J}$  dann auch suffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ .

Die Forderung der Existenz des Maßes  $\pi$  aus (4.5)(b) kann nicht fallengelassen werden.

Beispiel: Nimmt man als Experiment  $(R, \mathcal{P}(R))$  mit allen dis-

kreten normierten Maßen auf  $\mathcal{P}(\mathbb{R})$  und als Unter- $\sigma$ -Algebra  $\mathcal{S}_1$ , so sind die Dichten dieser normierten Maße bezüglich des zählenden Maßes zwar  $\mathcal{S}_1$ -meßbar,  $\mathcal{S}_1$  ist aber nicht suffizient für die Familie.

Das zählende Maß ist auf  $\mathcal{S}_1$  nicht lokalisierbar.

Ist das Experiment aus (4.5) durch ein  $\sigma$ -endliches Maß dominiert, so folgt daraus schon die Existenz des in (b) geforderten Maßes  $\kappa$ .

## § 5. Minimalsuffizienz

(5.1) Definition.  $(X, \mathcal{O}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  sei ein Experiment und  $\mathfrak{M}$  eine für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{O}$ . Dann heißt  $\mathfrak{M}$  minimalsuffizient für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ , wenn für jede für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  suffiziente Unter- $\sigma$ -Algebra  $\mathfrak{O}'$  von  $\mathcal{O}$  gilt  $\mathfrak{M} \subset \mathfrak{O}'$   $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ -fast. In diesem Fall wird auch das Unterexperiment  $(X, \mathfrak{M}, (P_{\mathfrak{J}|\mathfrak{M}})_{\mathfrak{J} \in \mathcal{O}})$  minimalsuffizient genannt.

Es soll jetzt die Existenz einer minimalsuffizienten Unter- $\sigma$ -Algebra  $\mathfrak{M}$  von  $\mathcal{O}$  für ein schwach dominiertes Experiment  $(X, \mathcal{O}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  bewiesen werden. Zunächst gilt:

(5.2) Satz.  $(X, \mathcal{O}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  sei ein Experiment und  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  durch ein  $\sigma$ -endliches Maß dominiert. Dann gibt es eine für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  minimalsuffiziente Unter- $\sigma$ -Algebra  $\mathfrak{M}$  von  $\mathcal{O}$ .

Zum Beweis vgl. [1], S. 439 Theorem 6.2 .

(5.3) Satz.  $(X, \mathcal{O}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  sei ein durch ein zu  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  äquivalentes Maß  $\mu$  schwach dominiertes Experiment. Dann gibt es eine für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  minimalsuffiziente Unter- $\sigma$ -Algebra  $\mathfrak{M}$  von  $\mathcal{O}$ .

Beweis.  $\mathfrak{O}'$  sei eine Unter- $\sigma$ -Algebra von  $\mathcal{O}$ . Notwendig für die Suffizienz von  $\mathfrak{O}'$  für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  ist die Suffizienz von  $\mathfrak{O}'$  für  $(P_{\mathfrak{J}}^{B_0})_{\mathfrak{J} \in \mathcal{O}}$ , wobei  $B_0$  ein festes Element von  $\mathfrak{O}'$  ist. Ist  $\mathfrak{O}'$  für  $(P_{\mathfrak{J}}^{B_0})_{\mathfrak{J} \in \mathcal{O}}$  suffizient, so ist auch die  $\sigma$ -Algebra  $\mathfrak{O}'_{B_0} := \mathfrak{O}' \cap B_0 \oplus \{X, \emptyset\} \cap (X \setminus B_0)$   
 $:= \{X_1 \cup X_2 : X_1 \in \mathfrak{O}' \cap B_0, X_2 \in \{X, \emptyset\} \cap (X \setminus B_0)\}$   
suffizient für  $(P_{\mathfrak{J}}^{B_0})_{\mathfrak{J} \in \mathcal{O}}$ .

Notwendig für die Minimalsuffizienz von  $\mathfrak{O}'$  für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$  ist die Minimalsuffizienz von  $\mathfrak{O}'$  für  $(P_{\mathfrak{J}}^{B_0})_{\mathfrak{J} \in \mathcal{O}}$ , wobei  $B_0$  ein

festes Element von  $\mathcal{G}$  ist.

In der Tat: Sei  $\mathcal{G}$  minimalstüffizient für  $(P_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$  und  $\mathcal{G}'$  minimalstüffizient für  $(P_{\mathcal{G}}^{B_0})_{\mathcal{G} \in \mathcal{O}}$ . Dann ist  $\mathcal{G}' \cap B_0 \subset \mathcal{G} \cap B_0$   $\mu$ -fast. Definiert man eine  $\sigma$ -Algebra

$$\mathcal{G}'' := \mathcal{G}' \cap B_0 \oplus \mathcal{G} \cap (X \setminus B_0),$$

so ist diese stüffizient für  $(P_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$ . Denn zu jedem  $A \in \mathcal{A}$  gibt es eine  $\mathcal{G}$ -meßbare Funktion  $E(1_A | \mathcal{G})$  und eine  $\mathcal{G}'$ -meßbare Funktion  $E(1_A | \mathcal{G}')$ , die ein bedingter  $(P_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{G}$  bzw. ein bedingter  $(P_{\mathcal{G}}^{B_0})_{\mathcal{G} \in \mathcal{O}}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{G}'$  ist. Also ist

$E(1_A | \mathcal{G}') \cdot 1_{B_0} + E(1_A | \mathcal{G}) \cdot 1_{X \setminus B_0}$  ein bedingter  $(P_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{G}''$ . Nun ist  $\mathcal{G}'' \subset \mathcal{G}$   $\mu$ -fast und daher  $\mathcal{G}'' = \mathcal{G}$   $\mu$ -fast, woraus folgt  $\mathcal{G}' \cap B_0 = \mathcal{G} \cap B_0$   $\mu$ -fast. Dies liefert die Minimalstüffizienz von  $\mathcal{G}$  für  $(P_{\mathcal{G}}^{B_0})_{\mathcal{G} \in \mathcal{O}}$ .

Es sei nun wieder  $T_{\mathcal{G}} := \{x \in X : \frac{dP_{\mathcal{G}}}{d\mu}(x) > 0\}$  für alle  $\mathcal{G} \in \mathcal{O}$ . Für jedes  $\mathcal{G}_0 \in \mathcal{O}$  ist  $(P_{\mathcal{G}_0}^{T_{\mathcal{G}_0}})_{\mathcal{G} \in \mathcal{O}}$  durch das  $\sigma$ -endliche Maß  $\mu^{T_{\mathcal{G}_0}}$  dominiert. Daher existiert für diese Familie eine minimalstüffiziente  $\sigma$ -Algebra  $\mathcal{G}'_{T_{\mathcal{G}_0}}$  nach (5.2), die o.E.d.A. alle  $\mu$ -Nullmengen aus  $\mathcal{A}$  enthalte.

$\mathcal{G}_{T_{\mathcal{G}_0}} := \mathcal{G}'_{T_{\mathcal{G}_0}} \cap T_{\mathcal{G}_0} \oplus \{X, \emptyset\} \cap (X \setminus T_{\mathcal{G}_0})$  ist dann ebenfalls eine minimalstüffiziente  $\sigma$ -Algebra für  $(P_{\mathcal{G}}^{T_{\mathcal{G}_0}})_{\mathcal{G} \in \mathcal{O}}$ .

Es seien jetzt  $\mathcal{G}_1, \mathcal{G}_2 \in \mathcal{O}$ . Es gilt

$$T_{\mathcal{G}_1} \cap T_{\mathcal{G}_2} = T_{\mathcal{G}_1} \cap \left\{ \frac{dP_{\mathcal{G}_2}}{d\mu} > 0 \right\} \text{ und}$$

$$T_{\mathcal{G}_1} \cap \left\{ \frac{dP_{\mathcal{G}_2}}{d\mu} > 0 \right\} = T_{\mathcal{G}_1} \cap \left\{ \frac{dP_{\mathcal{G}_2}^{T_{\mathcal{G}_1}}}{d\mu^{T_{\mathcal{G}_1}}} > 0 \right\} \mu\text{-fast.}$$

Nach (2.2) ist  $\left\{ \frac{dP_{\mathcal{G}_2}^{T_{\mathcal{G}_1}}}{d\mu^{T_{\mathcal{G}_1}}} > 0 \right\} \in \mathcal{G}'_{T_{\mathcal{G}_1}}$   $\mu^{T_{\mathcal{G}_1}}$ -fast. Da  $\mathcal{G}'_{T_{\mathcal{G}_1}}$  alle  $\mu$ -Nullmengen enthält, ist also

$$T_{\mathcal{G}_1} \cap \left\{ \frac{dP_{\mathcal{G}_2}^{T_{\mathcal{G}_1}}}{d\mu^{T_{\mathcal{G}_1}}} > 0 \right\} \in \mathcal{G}_{T_{\mathcal{G}_1}} \text{ und folglich auch } T_{\mathcal{G}_1} \cap T_{\mathcal{G}_2} \in \mathcal{G}_{T_{\mathcal{G}_1}}.$$

Analog erhält man  $T_{j_1} \cap T_{j_2} \in \mathcal{O}_{T_{j_2}}$ . Aus den Überlegungen am Anfang des Beweises folgt nun, wenn wir  $B := T_{j_1} \cap T_{j_2}$  setzen, daß  $\mathcal{O}_{T_{j_1}}$  und  $\mathcal{O}_{T_{j_2}}$  minimalstuffsizient für  $(P_j^B)_{j \in \Theta}$  sind. Weil außerdem  $\mathcal{O}'_{T_{j_1}}$  und  $\mathcal{O}'_{T_{j_2}}$  alle  $\mu$ -Nullmengen aus  $\mathcal{O}$  enthalten, gilt  $\mathcal{O}_{T_{j_1}} \cap B = \mathcal{O}_{T_{j_2}} \cap B$ . Das Mengensystem  $\bigcup_{j \in \Theta} (\mathcal{O}_{T_j} \cap T_j)$  ist somit durchschnittsstabil.

Notwendigerweise umfaßt jede für  $(P_j)_{j \in \Theta}$  suffiziente  $\sigma$ -Algebra, die alle  $\mu$ -Nullmengen aus  $\mathcal{O}$  enthält, dieses Mengensystem.  $\mathcal{A}$  sei die Menge aller Unter- $\sigma$ -Algebren von  $\mathcal{O}$ , die  $\bigcup_{j \in \Theta} (\mathcal{O}_{T_j} \cap T_j)$  und alle  $\mu$ -Nullmengen aus  $\mathcal{O}$  umfassen und für die  $\mu$  lokalisierbar bleibt.  $\mathcal{A}$  ist nicht leer, da  $\mathcal{O} \in \mathcal{A}$  ist. Der Durchschnitt aller  $\sigma$ -Algebren aus  $\mathcal{A}$  ist wieder eine  $\sigma$ -Algebra  $\mathcal{O}^*$ .

Wir behaupten:  $\mu|_{\mathcal{O}^*}$  ist lokalisierbar.

In der Tat: Sei  $(F_j)_{j \in J}$ ,  $J \neq \emptyset$ , eine Familie  $\mathcal{O}^*$ -meßbarer Mengen. Dann sind für jedes  $j \in J$   $F_j$  und alle Elemente aus  $\tilde{F}_j$  in jeder  $\sigma$ -Algebra aus  $\mathcal{A}$ . Daher ist jedes  $F \in \sup_{j \in J} \tilde{F}_j$  in jeder  $\sigma$ -Algebra aus  $\mathcal{A}$  enthalten, also auch in  $\mathcal{O}^*$ .

Da  $\bigcup_{j \in \Theta} (\mathcal{O}_{T_j} \cap T_j)$  durchschnittsstabil ist, gilt für jedes  $j_0 \in \Theta$

$$\begin{aligned} \mathcal{O}^* &= \mathcal{O}^* \cap T_{j_0} \oplus \mathcal{O}^* \cap (X \setminus T_{j_0}) \supset \mathcal{O}_{T_{j_0}} \cap T_{j_0} \oplus \mathcal{O}^* \cap (X \setminus T_{j_0}) \\ &\supset \bigcup_{j \in \Theta} (\mathcal{O}_{T_j} \cap T_j). \end{aligned}$$

$\mu$ , eingeschränkt auf die  $\sigma$ -Algebra

$\mathcal{O}_{T_{j_0}} \cap T_{j_0} \oplus \mathcal{O}^* \cap (X \setminus T_{j_0})$ , ist lokalisierbar.

Dies ergibt sich, da Lokalisierbarkeit nicht beim Übergang zu einem äquivalenten Maß verlorenggeht, wie folgt:

$\mu^{T_{j_0}}$  ist  $\sigma$ -endlich. Also gibt es ein dazu äquivalentes normiertes Maß  $\mu'$ .  $\mu'^{T_{j_0}}|_{\mathcal{O}_{T_{j_0}}}$  ist endlich, also lokalisierbar.

$\mu$  ist äquivalent zu  $\mu|_{T_{\vartheta_0}} + \mu|_{X \setminus T_{\vartheta_0}}$ , und  
 $(\mu|_{T_{\vartheta_0}} + \mu|_{X \setminus T_{\vartheta_0}}) | (\delta_{T_{\vartheta_0}} \cap T_{\vartheta_0} \oplus \delta^* \cap (X \setminus T_{\vartheta_0}))$   
 $= \mu|_{T_{\vartheta_0}} | \delta_{T_{\vartheta_0}} + \mu|_{X \setminus T_{\vartheta_0}} | \delta^*$  ist lokalisierbar.

Wegen der Minimalität von  $\delta^*$  folgt nun für alle  $\vartheta \in \mathcal{C}$

$$\delta^* \cap T_{\vartheta} = \delta_{T_{\vartheta}} \cap T_{\vartheta}.$$

Wie im Beweis von (2.3) gibt es eine Familie  $\delta^*$ -meßbarer Mengen  $(X_i)_{i \in I}$ ,  $I \neq \emptyset$ , so, daß jedes  $X_i$ ,  $i \in I$ , positives  $\sigma$ -endliches  $\mu$ -Maß hat,  $\mu(X_i \cap X_{i'}) = 0$  ist für alle  $i, i' \in I$  mit  $i \neq i'$ , aus  $\mu(A \cap X_i) = 0$ ,  $A \in \mathcal{A}$ , für alle  $i \in I$  folgt  $\mu(A) = 0$  und für jedes  $i_0 \in I$  ein  $\vartheta_0 \in \mathcal{C}$  existiert mit  $X_{i_0} \subset T_{\vartheta_0}$ . Den im dortigen Beweis verwendeten bedingten  $(P_{\vartheta})_{\vartheta \in \mathcal{C}}$ -Erwartungswert  $E(1_A | \delta)$ ,  $A \in \mathcal{A}$ , ersetzt man dabei durch  $E(1_A | \delta_{T_{\vartheta_0}})$ , den bedingten  $(P_{\vartheta}^{T_{\vartheta_0}})_{\vartheta \in \mathcal{C}}$ -Erwartungswert von  $1_A$  bezüglich  $\delta_{T_{\vartheta_0}}$ , wobei  $\vartheta_0 \in \mathcal{C}$  wie in (2.3) zu bestimmen ist.

Wie im Beweis von (2.4) kann man dann ein zu  $\mu$  äquivalentes Maß  $\nu$  angeben, das, auch auf  $\delta^*$  eingeschränkt, die endliche Teilmengeneigenschaft hat.

Für jedes  $E \in \delta^*$  mit  $\nu(E) < \infty$  gibt es eine abzählbare Teilmenge  $J \subset I$  mit  $E = \bigcup_{j \in J} E \cap X_j$   $\nu$ -fast. Zu jedem  $j \in J$  gibt es ein  $\vartheta_j \in \mathcal{C}$  mit  $E \cap X_j \in \delta^* \cap T_{\vartheta_j} = \delta_{T_{\vartheta_j}} \cap T_{\vartheta_j}$ . Da für jedes  $\vartheta_0 \in \mathcal{C}$  und  $A \in \mathcal{A}$  ein bedingter  $(P_{\vartheta}^{T_{\vartheta_0}})_{\vartheta \in \mathcal{C}}$ -Erwartungswert von  $1_A$  bezüglich  $\delta_{T_{\vartheta_0}}$   $E(1_A | \delta_{T_{\vartheta_0}})$  existiert, ist durch

$$\psi_E(1_A) := \sum_{j \in J} E(1_A | \delta_{T_{\vartheta_j}}) \cdot 1_{X_j \cap E},$$

wobei für jedes  $j \in J$   $E(1_A | \delta_{T_{\vartheta_j}})$  für alle  $x \in X$  nicht negativ gewählt sei, eine  $\delta^*$ -meßbare Funktion definiert mit  
 $\int_S \psi_E(1_A) dP_{\vartheta} = \int_S 1_A dP_{\vartheta}$  für alle  $\vartheta \in \mathcal{C}$  und  $S \in \delta^*$  mit  $S \subset E$ .

$(\overline{\nu_E(1_A)})^{\nu, \delta^*}_{E \in \delta^*, \nu(E) < \infty}$  ist daher ein Querschnitt, der wegen der Eigenschaften von  $\nu$  durch eine  $\delta^*$ -meßbare Funktion  $\nu(1_A)$  bestimmt ist, die ein bedingter  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ -Erwartungswert von  $1_A$  bezüglich  $\delta^*$  ist, da jedes  $T_{\mathcal{J}}, \mathcal{J} \in \mathcal{C}$ ,  $\sigma$ -endliches  $\nu|_{\delta^*}$ -Maß hat.

$\delta^*$  ist also suffizient und gemäß seiner Konstruktion auch minimalsuffizient für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ . Für  $\mathcal{M}$  kann man daher  $\delta^*$  nehmen.

(5.4) Korollar. Unter den Voraussetzungen des vorigen Satzes ist die kleinste alle  $\mu$ -Nullmengen von  $\mathcal{O}$  enthaltende Unter- $\sigma$ -Algebra von  $\mathcal{O}$ , für die  $\frac{dP_{\mathcal{J}}^{T_{\mathcal{J}'}}}{dP_{\mathcal{J}}} \cdot 1_{T_{\mathcal{J}'}}$  für alle  $\mathcal{J}, \mathcal{J}' \in \mathcal{C}$  meßbar ist und  $\mu$  lokalisierbar bleibt, eine minimalsuffiziente  $\sigma$ -Algebra für  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$ .

Beweis. Für jedes  $\mathcal{J}' \in \mathcal{C}$  ist  $(P_{\mathcal{J}}^{T_{\mathcal{J}'}})_{\mathcal{J} \in \mathcal{C}}$  durch  $P_{\mathcal{J}'}$  dominiert. Dann ist die für die Familie  $(P_{\mathcal{J}}^{T_{\mathcal{J}'}})_{\mathcal{J} \in \mathcal{C}}$  minimalsuffiziente  $\sigma$ -Algebra  $\delta'_{T_{\mathcal{J}'}}$  aus (5.3) als die kleinste alle  $P_{\mathcal{J}'}$ -Nullmengen von  $\mathcal{O}$  umfassende  $\sigma$ -Algebra, für die  $\frac{dP_{\mathcal{J}}^{T_{\mathcal{J}'}}}{dP_{\mathcal{J}'}}$  für alle  $\mathcal{J} \in \mathcal{C}$  meßbar ist, wählbar (vgl. den Beweis von (5.2) in [1]). Aus dem weiteren Vorgehen in (5.3) folgt dann die Behauptung.

## II. Schwach suffiziente Experimente

### § 6. Schwache Suffizienz beim Vergleich zweier Experimente

Im folgenden seien  $(X, \mathcal{A}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  und  $(Y, \mathcal{I}, (Q_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  Experimente.

(6.1) Definition. Eine Abbildung  $\varphi: \mathcal{I} \times X \rightarrow \mathbb{R}$  heit schwacher Markovkern von  $(X, \mathcal{A}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  nach  $(Y, \mathcal{I})$ , wenn

- (1)  $0 \leq \varphi(C, \cdot)$   $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ -fast berall fr alle  $C \in \mathcal{I}$  ist,
- (2)  $\varphi(\bigcup_{n=1}^{\infty} C_n, \cdot) = \sum_{n=1}^{\infty} \varphi(C_n, \cdot)$   $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ -fast berall fr jede Folge paarweiser disjunkter Mengen  $(C_n)_{n \in \mathbb{N}}$ ,  $C_n \in \mathcal{I}$ , ist,
- (3)  $\varphi(Y, \cdot) = 1$   $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ -fast berall gilt,
- (4)  $\varphi(C, \cdot)$  fr alle  $C \in \mathcal{I}$   $\mathcal{A}$ -mebar ist.

Zusatz: Gelten die Beziehungen (1) bis (4) nicht nur  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}}$ -fast berall, sondern fr alle  $x \in X$ , so heit  $\varphi$  ein Markovkern.

Die Festsetzung  $P_{\mathfrak{J}}^*(C) := \int \varphi(C, x) P_{\mathfrak{J}}(dx)$  fr alle  $C \in \mathcal{I}$  liefert fr jedes  $\mathfrak{J} \in \mathcal{O}$  ein normiertes Ma auf  $\mathcal{I}$ . Dies folgt unmittelbar aus (6.1). Insbesondere die  $\sigma$ -Additivitt ergibt sich aus Bedingung (2), da das Integral einer Summe fast berall nicht negativer mebarer Funktionen gleich der Summe der Integrale der einzelnen Summanden ist (vgl. [2], Korollar 11.5). Wir erhalten also ein neues Experiment  $(Y, \mathcal{I}, (P_{\mathfrak{J}}^*)_{\mathfrak{J} \in \mathcal{O}})$ . Insbesondere sagen wir:

(6.2) Definition. Ein schwacher Markovkern von  $(X, \mathcal{A}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  nach  $(Y, \mathcal{I})$  heit schwacher Markovkern von  $(X, \mathcal{A}, (P_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$  nach  $(Y, \mathcal{I}, (Q_{\mathfrak{J}})_{\mathfrak{J} \in \mathcal{O}})$ , wenn  $P_{\mathfrak{J}}^* = Q_{\mathfrak{J}}$  fr alle  $\mathfrak{J} \in \mathcal{O}$  ist.

(6.3) Beispiel. Es sei  $X = Y$ ,  $\mathcal{I}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{A}$  und  $Q_{\mathfrak{J}} = P_{\mathfrak{J}}|_{\mathcal{I}}$ . Dann liefert  $\varphi(C, x) := 1_C(x)$  einen schwa-

chen Markovkern von  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$  nach  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$ . Ist außerdem  $\mathcal{I}$  suffizient für  $(P_y)_{y \in \mathcal{C}}$  und für jedes  $A \in \mathcal{A}$   $E(1_A | \mathcal{I})$  eine Version des bedingten  $(P_y)_{y \in \mathcal{C}}$ -Erwartungswertes, so liefert  $\psi(A, x) := E(1_A | \mathcal{I})(x)$  einen schwachen Markovkern von  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$  nach  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$ .

(6.3) zeigt, daß man den Begriff der Suffizienz verallgemeinern kann, wenn zwei beliebige Experimente mit gleicher Parametermenge gegeben sind, indem man die Existenz eines schwachen Markovkerns vom ersten nach dem zweiten Experiment und eines weiteren schwachen Markovkerns in umgekehrter Richtung fordert. Dies führt in Anlehnung an Morse und Sacksteder [12] zur Definition isomorpher Experimente.

(6.4) Definition. Zwei Experimente  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$  und  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$  heißen isomorph, wenn es einen schwachen Markovkern von  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$  nach  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$  gibt und einen anderen von  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$  nach  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$ .

Der Begriff der Suffizienz läßt sich in noch etwas anderer, weniger weitgehender Weise verallgemeinern. Die folgende Betrachtung dient der Motivation des weiteren Vorgehens. Sie folgt im wesentlichen einer Idee von Csizsár [6].

Gegeben sei ein schwacher Markovkern  $\varphi$  von  $(X, \mathcal{A}, (P_y)_{y \in \mathcal{C}})$  nach  $(Y, \mathcal{I}, (Q_y)_{y \in \mathcal{C}})$  (bei Csizsár ist  $\varphi$  ein Markovkern). Die beiden Experimente lassen sich als Unterexperimente eines weiteren Experimentes auffassen, das nun konstruiert werden soll.

(6.5) Lemma.  $\mu$  sei ein bezüglich  $(P_y)_{y \in \mathcal{C}}$  absolut stetiges Maß auf  $\mathcal{A}$ . Die Festsetzung

$$\mu'(A \times C) := \int_A \varphi(C, x) \mu(dx) \quad \text{für alle } A \in \mathcal{A} \text{ und } C \in \mathcal{I}$$

liefert dann ein Maß auf  $(X \times Y, \mathcal{O} \otimes \mathcal{I})$ .

Beweis. Die angegebene Definition ist sinnvoll, da  $\varphi(C, \cdot)$   $(P_\delta)_{\delta \in \mathcal{C}}$ -fast überall und folglich auch  $\mu$ -fast überall nicht negativ ist.

Es sei  $\emptyset \neq D \times E \in \mathcal{O} \times \mathcal{I}$  und  $(A_n \times C_n)_{n \in \mathbb{N}}$  eine Folge paarweiser disjunkter Mengen aus  $\mathcal{O} \times \mathcal{I}$  mit  $D \times E = \bigcup_{n \in \mathbb{N}} A_n \times C_n$ . Für jedes  $x_0 \in D$  sei  $J_{x_0} := \{n \in \mathbb{N} : x_0 \in A_n\}$ .  $J_{x_0}$  ist nicht leer. Es gilt  $E = \sum_{n \in J_{x_0}} C_n$  für jedes  $x_0 \in D$ , da für alle  $x_0 \in D$  gilt  $\{x_0\} \times E = \bigcup_{n \in J_{x_0}} \{x_0\} \times C_n$  und für  $n \in J_{x_0}$  die Mengen  $C_n$  paarweise disjunkt sein müssen. Weiter gilt für  $\mu$ -fast alle  $x_0 \in D$  die Gleichungskette

$$\begin{aligned} \sum_{n \in \mathbb{N}} 1_{A_n}(x_0) \cdot \varphi_+(C_n, x_0) &= \sum_{n \in J_{x_0}} \varphi_+(C_n, x_0) = \varphi_+(\bigcup_{n \in J_{x_0}} C_n, x_0) \\ &= \varphi_+(E, x_0) = 1_D(x_0) \cdot \varphi_+(E, x_0), \end{aligned}$$

wobei für  $C \in \mathcal{I}$   $\varphi_+(C, \cdot)$  der Positivteil der  $\mathcal{O}$ -meßbaren Funktion  $\varphi(C, \cdot)$  sei. Da das erste und letzte Glied für  $x_0 \in X \setminus D$  Null sind, hat man für  $\mu$ -fast alle  $x \in X$

$$\begin{aligned} 1_D(x) \cdot \varphi_+(E, x) &= \sum_{n \in \mathbb{N}} 1_{A_n}(x) \cdot \varphi_+(C_n, x) \quad \text{und folglich} \\ \mu'(D \times E) &= \int_D \varphi(E, x) \mu(dx) = \int 1_D(x) \cdot \varphi_+(E, x) \mu(dx) \\ &= \int \left( \sum_{n \in \mathbb{N}} 1_{A_n}(x) \cdot \varphi_+(C_n, x) \right) \mu(dx) = \sum_{n \in \mathbb{N}} \int 1_{A_n}(x) \cdot \varphi_+(C_n, x) \mu(dx) \\ &= \sum_{n \in \mathbb{N}} \int_{A_n} \varphi(C_n, x) \mu(dx) = \sum_{n \in \mathbb{N}} \mu'(A_n \times C_n) \end{aligned}$$

$\mu'$  ist also eine monotone,  $\sigma$ -additive Funktion mit

$\mu'(\emptyset) = \mu'(\emptyset \times \emptyset) = 0$  auf dem Semiring  $\mathcal{O} \times \mathcal{I}$ . Nach [15], chapter 2 § 7, läßt sich  $\mu'$  zu einem Maß auf einer  $\mathcal{O} \times \mathcal{I}$  umfassenden  $\sigma$ -Algebra fortsetzen. Insbesondere hat man also eine Fortsetzung von  $\mu'$  zu einem Maß auf  $\mathcal{O} \otimes \mathcal{I}$ .

Für alle  $\delta \in \mathcal{C}$  sei  $P_\delta$  ein normiertes Maß  $P'_\delta$  auf  $\mathcal{O} \otimes \mathcal{I}$  zu-

geordnet, indem man in (6.5)  $\mu$  durch  $P_{\mathcal{G}}$  und  $\mu'$  durch  $P'_{\mathcal{G}}$  ersetzt. Damit erhalten wir ein neues Experiment  $(X \times Y, \alpha \otimes \mathcal{I}, (P'_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}})$ .

$\mathcal{X}' := \{X \times C : C \in \mathcal{I}\}$  und  $\alpha' := \{A \times Y : A \in \alpha\}$

sind Unter- $\sigma$ -Algebren von  $\alpha \otimes \mathcal{I}$ . Sie liefern Unterexperimente

$(X \times Y, \alpha', (P'_{\mathcal{G}}|_{\alpha'})_{\mathcal{G} \in \mathcal{O}})$  und  $(X \times Y, \mathcal{X}', (P'_{\mathcal{G}}|_{\mathcal{X}'} )_{\mathcal{G} \in \mathcal{O}})$ .

Vermöge der Abbildungen

$T_1: \alpha \ni A \mapsto A \times Y \in \alpha'$  und  $T_2: \mathcal{I} \ni C \mapsto X \times C \in \mathcal{X}'$

kann man wegen  $P'_{\mathcal{G}}(T_1(A)) = P_{\mathcal{G}}(A)$  und  $P'_{\mathcal{G}}(T_2(C)) = Q_{\mathcal{G}}(C)$  diese beiden Unterexperimente mit den Experimenten

$(X, \alpha, (P_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}})$  bzw.  $(Y, \mathcal{I}, (Q_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}})$  identifizieren.

$pr_1: X \times Y \ni (x, y) \mapsto x \in X$  und  $pr_2: X \times Y \ni (x, y) \mapsto y \in Y$

seien die Projektionsabbildungen von  $X \times Y$  auf  $X$  bzw.  $Y$ .

Es gilt  $P_{\mathcal{G}} = P'_{\mathcal{G}} \circ pr_1^{-1}$  und  $Q_{\mathcal{G}} = P'_{\mathcal{G}} \circ pr_2^{-1}$  für alle  $\mathcal{G} \in \mathcal{O}$ .

Die Menge der  $\alpha'$ -meßbaren Funktionen ist gleich der Menge

der Funktionen der Gestalt  $g \circ pr_1$ , wobei  $g$   $\alpha$ -meßbar ist;

die Menge der  $\mathcal{X}'$ -meßbaren Funktionen ist gleich der Menge

der Funktionen der Gestalt  $h \circ pr_2$ , wobei  $h$   $\mathcal{I}$ -meßbar ist.

(6.6) Lemma.  $\alpha'$  ist suffizient für  $(P'_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$ .

Beweis. Für jedes  $A \times C \in \alpha \otimes \mathcal{I}$  ist  $(1_A \cdot \varphi(C, \cdot)) \circ pr_1$  eine Version eines bedingten  $(P'_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}$ -Erwartungswertes von  $1_{A \times C}$

bezüglich  $\alpha'$ ; denn die  $\alpha'$ -Meßbarkeit ist klar und weiter gilt

$$\begin{aligned} \int_{A' \times Y} 1_{A \times C} dP'_{\mathcal{G}} &= \int_A 1_A \cdot \varphi(C, \cdot) dP_{\mathcal{G}} = \int_A 1_A \cdot \varphi(C, \cdot) d(P'_{\mathcal{G}} \circ pr_1^{-1}) \\ &= \int_{A' \times Y} (1_A \cdot \varphi(C, \cdot)) \circ pr_1 dP'_{\mathcal{G}} \text{ für alle } A' \times Y \in \alpha' \text{ und } \mathcal{G} \in \mathcal{O}. \end{aligned}$$

Nun ist die Menge

$\{F \in \alpha \otimes \mathcal{I} : \text{Zu } F \text{ gibt es eine } (P'_{\mathcal{G}})_{\mathcal{G} \in \mathcal{O}}\text{-fast überall nicht negative } \alpha'\text{-meßbare Funktion } f_F \text{ mit } \int_{A \times Y} 1_F dP'_{\mathcal{G}} = \int_{A \times Y} f_F dP'_{\mathcal{G}} \text{ für alle } A \times Y \in \alpha' \text{ und } \mathcal{G} \in \mathcal{O}\}$

offensichtlich ein Dynkin-System, das gleich der  $\sigma$ -Algebra  $\mathcal{Q} \otimes \mathcal{X}$  ist, da  $\mathcal{Q} \times \mathcal{X}$  ein durchschnittsstabiler Erzeuger von beiden ist (vgl. [2], S.18 Satz 2.3).

Das Lemma (6.6) macht es plausibel  $(Y, \mathcal{X}, (Q_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$  in einem verallgemeinerten Sinn als suffizient für  $(X, \mathcal{Q}, (P_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$  zu betrachten, wenn  $\mathcal{X}'$  suffizient für  $(P'_{\mathcal{Y}})_{\mathcal{Y} \in \Theta}$  ist. Um dies nur mit Hilfe der Ausgangsexperimente definieren zu können, beweisen wir das folgende

(6.7) Lemma.  $\mathcal{X}'$  ist genau dann suffizient für  $(P'_{\mathcal{Y}})_{\mathcal{Y} \in \Theta}$ , wenn es zu jedem  $A \in \mathcal{Q}$  eine  $\mathcal{X}$ -meßbare Funktion  $f_A$  gibt mit

$$\int_C f_A(y) Q_{\mathcal{Y}}(dy) = \int_C \varphi(C, x) P_{\mathcal{Y}}(dx) \quad \text{für alle } C \in \mathcal{X} \text{ und } \mathcal{Y} \in \Theta.$$

Beweis. Wenn es für jedes  $A \in \mathcal{Q}$  solch ein  $f_A$  gibt, liefert  $\nu(A, y) := f_A(y)$  einen schwachen Markovkern von  $(Y, \mathcal{X}, (Q_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$  nach  $(X, \mathcal{Q}, (P_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$ . (6.6) zeigt dann die Suffizienz von  $\mathcal{X}'$  für  $(P'_{\mathcal{Y}})_{\mathcal{Y} \in \Theta}$ , indem man dort  $P_{\mathcal{Y}}$  mit  $Q_{\mathcal{Y}}$  und  $\mathcal{Q}$  mit  $\mathcal{X}$  vertauscht.

Wenn  $\mathcal{X}'$  suffizient für  $(P'_{\mathcal{Y}})_{\mathcal{Y} \in \Theta}$  ist, so läßt sich für jede Indikatorfunktion  $1_{A \times Y}$ ,  $A \in \mathcal{Q}$ , eine Version des bedingten  $(P'_{\mathcal{Y}})_{\mathcal{Y} \in \Theta}$ -Erwartungswertes davon bezüglich  $\mathcal{X}'$  wählen, die die Gestalt  $f_{A \times Y} \circ \text{pr}_2$  hat, wobei  $f_{A \times Y}$   $\mathcal{X}$ -meßbar ist. Es gilt

$$\begin{aligned} \int_A \varphi(C, x) P_{\mathcal{Y}}(dx) &= \int_{X \times C} 1_{A \times Y}(x, y) P'_{\mathcal{Y}}(dx, dy) \\ &= \int_{X \times C} f_{A \times Y} \circ \text{pr}_2(x, y) P'_{\mathcal{Y}}(dx, dy) = \int_C f_{A \times Y}(y) P'_{\mathcal{Y}} \circ \text{pr}_2^{-1}(dy) \\ &= \int_C f_{A \times Y}(y) Q_{\mathcal{Y}}(dy) \quad \text{für alle } C \in \mathcal{X} \text{ und } \mathcal{Y} \in \Theta. \end{aligned}$$

Für jedes  $A \in \mathcal{Q}$  leistet  $f_{A \times Y}$  also das Gewünschte.

Die Beweise von (6.6) und (6.7) gehen auf Csiszár [6] zurück. Die folgende Definition gibt nun die angekündigte

Verallgemeinerung des Begriffes Suffizienz.

(6.8) Definition. Es existiere ein schwacher Markovkern  $\varphi$  von  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  nach  $(Y, \mathcal{I})$ .

(1)  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  heit schwach suffizient fr

$(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$ , wenn es einen schwachen Markovkern  $\psi$  von  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  nach  $(X, \mathcal{A})$  gibt mit

$$\int_C \psi(A, y) Q_{\vartheta}(dy) = \int_A \varphi(C, x) P_{\vartheta}(dx) \quad \text{fr alle } \vartheta \in \Theta, A \in \mathcal{A} \text{ und } C \in \mathcal{I}.$$

(2)  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  heit paarweise schwach suffizient fr

$(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$ , wenn es fr jede Wahl von  $\vartheta_1, \vartheta_2 \in \Theta$  einen schwachen Markovkern  $\psi_{\vartheta_1, \vartheta_2}$  von  $(Y, \mathcal{I}, (Q_{\vartheta_1}, Q_{\vartheta_2}))$  nach  $(X, \mathcal{A})$  gibt mit

$$\int_C \psi_{\vartheta_1, \vartheta_2}(A, y) Q_{\vartheta_i}(dy) = \int_A \varphi(C, x) P_{\vartheta_i}(dx) \quad \text{fr alle } A \in \mathcal{A}, C \in \mathcal{I} \text{ und fr } i = 1, 2.$$

Aus der Definition folgt unmittelbar, da  $\varphi$  bzw.  $\psi$  schwache Markovkerne von  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  nach  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  bzw. von  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  nach  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  sind, indem man  $A = X$  bzw.  $C = Y$  setzt.

Wenn  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$  schwach suffizient fr  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  ist, so ist auch  $(X, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  schwach suffizient fr  $(Y, \mathcal{I}, (Q_{\vartheta})_{\vartheta \in \Theta})$ .

Das Beispiel (6.3) zeigt, da jedes suffiziente Unterexperiment schwach suffizient fr das Ausgangsexperiment ist und da ein Experiment fr jedes seiner suffizienten Unterexperimente schwach suffizient ist.

Sacksteder [14] zeigt, da zwei Experimente isomorph sind, wenn ein Experiment davon fr das andere schwach suffizient

ist, was hier unmittelbar aus der Definitionen folgt, und bei isomorphen Experimenten das eine für das andere paarweise schwach suffizient ist. Er gibt Beispiele an, die zeigen, daß die Umkehrungen dieser beiden Implikationen im allgemeinen nicht gelten. Er beweist aber, daß

$(Y, \mathcal{I}, (Q_{\mathcal{V}})_{\mathcal{V} \in \Theta})$  für  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}})$  schwach suffizient ist, wenn  $(Y, \mathcal{I}, (Q_{\mathcal{V}})_{\mathcal{V} \in \Theta})$  für  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}})$  paarweise schwach suffizient ist und  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}}$  durch ein  $\sigma$ -endliches Maß dominiert ist. Die Verallgemeinerung davon für den schwach dominierten Fall geben wir im nächsten Paragraphen.

Im letzten Teil dieses Paragraphen wollen wir eine Beziehung zwischen schwach suffizienten Unterexperimenten und suffizienten Unterexperimenten behandeln. Wir benutzen folgende Bezeichnungen:

$(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}})$  sei ein Experiment und  $(P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}}$  sei durch ein dazu äquivalentes Maß  $\mu$  schwach dominiert. Es sei für alle  $\mathcal{V} \in \Theta$   $T_{\mathcal{V}} := \{x \in X : \frac{dP_{\mathcal{V}}}{d\mu}(x) > 0\}$ .  $(X, \mathcal{I}, (P_{\mathcal{V}}|_{\mathcal{I}})_{\mathcal{V} \in \mathcal{C}})$  sei ein für  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{C}})$  schwach suffizientes Unterexperiment. Es gibt also schwache Markovkerne  $\varphi$  auf  $X \times \mathcal{I}$  und  $\psi$  auf  $X \times \mathcal{A}$  mit  $\int_A \varphi(C, \cdot) dP_{\mathcal{V}} = \int_C \psi(A, \cdot) dP_{\mathcal{V}}$  für alle  $A \in \mathcal{A}$ ,  $C \in \mathcal{I}$  und  $\mathcal{V} \in \Theta$ . Schließlich sei  $\mathcal{F} := \{A \in \mathcal{A} : \psi(A, \cdot) = 1_A \mu\text{-fast überall}\}$ .

(6.9) Lemma. (a) Es ist  $\mathcal{F} \subset \mathcal{I}$   $\mu$ -fast.

(b)  $\mathcal{F}$  ist eine  $\sigma$ -Algebra, die  $T_{\mathcal{V}}$  für alle  $\mathcal{V} \in \Theta$  enthält.

Beweis. (a) ist klar.

(b): Zunächst hat man  $X \in \mathcal{F}$ ; denn  $\psi(X, \cdot) = 1 = 1_X \mu$ -fast überall. Sei  $S \in \mathcal{F}$ , so gilt  $\psi(X \setminus S, \cdot) = \psi(X, \cdot) - \psi(S, \cdot) = 1 - 1_S = 1_{X \setminus S} \mu$ -fast überall, und daher ist  $X \setminus S \in \mathcal{F}$ .

Sei  $(S_n)_{n \in \mathbb{N}}$  eine Folge von Mengen aus  $\mathcal{O}$ , so ist

$$\psi\left(\bigcup_{n=1}^{\infty} S_n, \cdot\right) \leq \sum_{n=1}^{\infty} 1_{S_n} \quad \mu\text{-fast überall, und da}$$

$$\psi\left(\bigcup_{n=1}^{\infty} S_n, \cdot\right) \leq 1 \quad \mu\text{-fast überall ist, folgt}$$

$$\psi\left(\bigcup_{n=1}^{\infty} S_n, \cdot\right) \leq 1_{\bigcup_{n=1}^{\infty} S_n} \quad \mu\text{-fast überall. Andererseits ist}$$

$$1_{S_n} = \psi(S_n, \cdot) \leq \psi\left(\bigcup_{n=1}^{\infty} S_n, \cdot\right) \quad \mu\text{-fast überall für alle } n \in \mathbb{N},$$

$$\text{so daß man } \psi\left(\bigcup_{n=1}^{\infty} S_n, \cdot\right) = 1_{\bigcup_{n=1}^{\infty} S_n} \text{ erhält.}$$

Es ist also  $\bigcup_{n=1}^{\infty} S_n \in \mathcal{O}$  und  $\mathcal{O}$  somit eine  $\sigma$ -Algebra.

Für  $\mathcal{V}_0 \in \mathcal{O}$  ist

$$\int \psi(T_{\mathcal{V}_0}, \cdot) \cdot 1_{T_{\mathcal{V}_0}} \, dP_{\mathcal{V}_0} = \int \psi(T_{\mathcal{V}_0}, \cdot) \, dP_{\mathcal{V}_0} = P_{\mathcal{V}_0}(T_{\mathcal{V}_0}) = 1$$

Da  $0 \leq \psi(T_{\mathcal{V}_0}, \cdot) \leq 1$   $\mu$ -fast überall ist, folgt

$$\psi(T_{\mathcal{V}_0}, \cdot) \cdot 1_{T_{\mathcal{V}_0}} \leq 1_{T_{\mathcal{V}_0}} \quad \mu\text{-fast überall und folglich}$$

$$1_{T_{\mathcal{V}_0}} = \psi(T_{\mathcal{V}_0}, \cdot) \cdot 1_{T_{\mathcal{V}_0}} \quad \mu\text{-fast überall, was zusammen mit}$$

$0 \leq \psi(T_{\mathcal{V}_0}, \cdot)$  zu  $\psi(T_{\mathcal{V}_0}, \cdot) \geq 1_{T_{\mathcal{V}_0}}$   $\mu$ -fast überall führt.

Weiter ist wegen  $\varphi(X, \cdot) = 1$   $\mu$ -fast überall

$$1 \geq \int 1_{T_{\mathcal{V}_0}} \, dP_{\mathcal{V}_0} = \int \psi(T_{\mathcal{V}_0}, \cdot) \, dP_{\mathcal{V}_0} \geq 0 \quad \text{für alle } \mathcal{V} \in \mathcal{O}.$$

Also ist  $\psi(T_{\mathcal{V}}, \cdot) = 1_{T_{\mathcal{V}}}$   $\mu$ -fast überall für alle  $\mathcal{V} \in \mathcal{O}$ .

(6.10) Lemma. (a) Für jedes  $f \in \mathcal{L}^{\infty}(\mathcal{A}, \mu)$  gibt es eine Funk-

tion  $\psi(f) \in \mathcal{L}^{\infty}(\mathcal{X}, \mu)$  ( $\psi$   $\mu$ -fast überall linear in  $f$ ) mit

$$\|\psi(f)\|_{\infty} \leq \|f\|_{\infty} \quad \text{und} \quad \int \varphi(C, \cdot) f \, dP_{\mathcal{V}} = \int_C \psi(f) \, dP_{\mathcal{V}}$$

für alle  $C \in \mathcal{X}$  und  $\mathcal{V} \in \mathcal{O}$ .

(b) Für jedes  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$  gibt es ein  $\psi'(g) \in \mathcal{L}^1(\mathcal{A}, \mu)$  mit

$$\int \psi(f) \cdot g \, d\mu = \int f \cdot \psi'(g) \, d\mu \quad \text{für alle } f \in \mathcal{L}^{\infty}(\mathcal{A}, \mu).$$

Beweis. (a): Sei  $f = \sum_{v=1}^n c_v \cdot 1_{A_v}$ , wobei  $c_v \in \mathbb{R}$  und  $A_v \in \mathcal{O}$

paarweise disjunkt mit  $\mu(A_v) > 0$  sind für  $v = 1, \dots, n$ , dann

leistet wegen  $\|f\|_{\infty} = \sup_v |c_v| \geq \sup_v |c_v| \|\psi(1_{A_v}, \cdot)\|_{\infty} \geq \|\psi(f)\|_{\infty}$

die Funktion  $\psi(f) = \sum_{v=1}^n c_v \cdot \psi(1_{A_v}, \cdot)$  das Gewünschte.

Wenn  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  das Supremum einer isotonen Folge  $(f_n)_{n \in \mathbb{N}}$  von überall nicht negativen Funktionen des eben behandelten Typs ist, erfüllt nach der Wahl der  $\psi(f_n)$ ,  $n \in \mathbb{N}$ , die stets  $\mu$ -fast überall eine isotone Folge bilden,  $\sup_{n \in \mathbb{N}} \psi(f_n)$  die geforderte Bedingung. Für beliebiges  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  kann man als  $\psi(f)$   $\psi(f_+) - \psi(f_-)$  nehmen.

(b): Sei jetzt  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$ . Dann hat

$T_g := \{x \in X : g(x) \neq 0\}$   $\sigma$ -endliches  $\mu$ -Maß. Deswegen gibt es eine abzählbare Teilmenge  $\mathcal{O}_g \subset \mathcal{O}$  mit  $T_g \subset \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}$   $\mu$ -fast. Man vergleiche dazu (2.3) und (2.4). Für alle  $A \in \mathcal{A}$  mit  $A \subset X \setminus \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}$  gilt zufolge Teil (a) und (6.9)

$0 \leq \psi(1_A) = \psi(A, \cdot) \leq \psi(X \setminus \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}, \cdot) = 1_{X \setminus \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}}$   $\mu$ -fast überall. Daher ist  $\psi(1_A) \cdot 1_{\bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}} = 0$   $\mu$ -fast überall für alle  $A \subset X \setminus \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}$ . Weiter ist

$$\int \psi(1_A) \cdot g \, d\mu = \int \psi(1_A) \cdot g \cdot 1_{\bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}} \, d\mu \text{ für alle } A \in \mathcal{A}.$$

Für  $g \in \mathcal{L}_+^1(\mathcal{A}, \mu)$  liefert daher die Zuordnung

$A \mapsto \int \psi(1_A) \cdot g \, d\mu$  für alle  $A \in \mathcal{A}$  ein endliches Maß auf  $\mathcal{A}$ , das bezüglich  $\mu^T$ ,  $T := \bigcup_{\mathcal{O} \in \mathcal{O}_g} T_{\mathcal{O}}$ , absolut stetig ist. Unter Benutzung des Satzes von Radon-Nikodym erhält man daher eine Funktion  $\psi'(g) \in \mathcal{L}_+^1(\mathcal{A}, \mu)$  mit

$$\int \psi(1_A) \cdot g \, d\mu = \int 1_A \cdot \psi'(g) \, d\mu \text{ für alle } A \in \mathcal{A}. \text{ Für ein beliebiges } g \in \mathcal{L}^1(\mathcal{A}, \mu) \text{ erfüllt } \psi'(g) = \psi'(g_+) - \psi'(g_-) \text{ die Gleichung } \int \psi(1_A) \cdot g \, d\mu = \int 1_A \cdot \psi'(g) \, d\mu \text{ für alle } A \in \mathcal{A}.$$

In der gleichen Weise, wie die Existenz eines  $\psi(f)$  für alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  in (a) nachgewiesen wurde, zeigt man, daß

$$\int \psi(f) \cdot g \, d\mu = \int f \cdot \psi'(g) \, d\mu \text{ ist für alle } f \in \mathcal{L}^\infty(\mathcal{A}, \mu) \text{ und}$$

$g \in \mathcal{L}^1(\mathcal{A}, \mu)$ .

Aus den Gleichungen, die  $\psi(f)$ ,  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$ , und  $\psi'(g)$ ,  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$ , erfüllen, folgt, daß diese Funktionen  $\mu$ -fast überall eindeutig bestimmt sind. Daher liefern die Zuordnungen  $\tilde{f} \mapsto \widetilde{\psi(f)}$  für  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  bzw.  $\tilde{g} \mapsto \widetilde{\psi'(g)}$  für  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$  lineare, positive und normierte Abbildungen von  $L^\infty(\mathcal{A}, \mu)$  in sich selbst bzw. von  $L^1(\mathcal{A}, \mu)$  in sich selbst.

(6.11) Satz.  $\mathcal{G}$  ist eine suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{A}$  für  $(P_A)_{A \in \mathcal{G}}$ .

Beweis. Für jedes  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  sei  $\psi^1(f) \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  mit  $\psi^1(f) = \psi(f)$   $\mu$ -fast überall und  $\psi^n(f) \in \mathcal{L}^\infty(\mathcal{A}, \mu)$ ,  $n > 1$ , mit  $\psi^n(f) = \psi(\psi^{n-1}(f))$   $\mu$ -fast überall. Schließlich sei  $\psi_n(f) \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  mit  $\psi_n(f) = \frac{1}{n} \sum_{v=1}^n \psi^v(f)$   $\mu$ -fast überall,  $n \in \mathbb{N}$ .

Aus  $\|\psi(f)\|_\infty \leq \|f\|_\infty$  folgt  $\|\psi^n(f)\|_\infty \leq \|f\|_\infty$  und daraus mit der Dreiecksungleichung  $\|\psi_n(f)\|_\infty \leq \|f\|_\infty$  für alle  $n \in \mathbb{N}$ , wobei  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  ist.

Die Zuordnung  $\tilde{f} \mapsto \widetilde{\psi_n(f)}$  für alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  liefert für jedes  $n \in \mathbb{N}$  eine positive, normierte lineare Abbildung von  $L^\infty(\mathcal{A}, \mu)$  in sich.

Nach [8], 3.3.1, ist die Menge der positiven, normierten linearen Abbildungen kompakt bezüglich der größten Topologie, für die die Zuordnungen  $l \mapsto \int l(f) \cdot g \, d\mu$  stetige Abbildungen liefern, sofern  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$ ,  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$  und  $l$  eine positive, normierte lineare Abbildung von  $L^\infty(\mathcal{A}, \mu)$  in sich ist und  $l(f)$  in  $l(\tilde{f})$  liegt. (In [8] ist  $\mu$   $\sigma$ -endlich vorausgesetzt. Es wird aber nur die schwache Kompaktheit der Einheitskugel von  $L^\infty(\mathcal{A}, \mu)$  gebraucht, was aus der auch

für lokalisierbares  $\mu$  mit der endlichen Teilmengeneigenschaft geltenden Gleichheit  $L^\infty(\mathcal{A}, \mu) = L^1(\mathcal{A}, \mu)$  (vgl. [15], S. 364 Theorem 4) folgt.)

Daher gibt es eine positive, normierte lineare Abbildung  $l$  von  $L^\infty(\mathcal{A}, \mu)$  in sich, die Häufungspunkt der durch die  $\psi_n$ ,  $n \in \mathbb{N}$ , definierten Abbildungen ist. Für jedes  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  sei  $l(f)$  ein Element aus  $L^1(\mathcal{A})$ .

Da für jedes  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  und  $n \in \mathbb{N}$   $\mu$ -fast überall  $\psi_n(f) - \psi(\psi_n(f)) = \frac{1}{n} (\psi(f) - \psi^{n+1}(f))$  ist, gilt  $\|\psi_n(f) - \psi(\psi_n(f))\|_\infty \leq \frac{2}{n} \|f\|_\infty$ . Sei nun  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$ , dann gibt es für jedes  $\varepsilon > 0$  ein  $n_0$  so, daß

$$\left| \int (\psi_n(f) - \psi(\psi_n(f))) \cdot g \, d\mu \right| < \frac{\varepsilon}{5} \quad \text{ist für alle } n \geq n_0.$$

Weiter gibt es ein  $n_1 \geq n_0$  so, daß

$$\left| \int (l(f) - \psi_{n_1}(f)) \cdot g \, d\mu \right| < \frac{\varepsilon}{5} \quad \text{und}$$

$$\left| \int (l(f) - \psi_{n_1}(f)) \cdot \psi'(g) \, d\mu \right| < \frac{\varepsilon}{5} \quad \text{ist. Damit folgt}$$

$$\left| \int (l(f) - \psi(l(f))) \cdot g \, d\mu \right| \leq \left| \int (l(f) - \psi_{n_1}(f)) \cdot g \, d\mu \right|$$

$$+ \left| \int (\psi_{n_1}(f) - \psi(\psi_{n_1}(f))) \cdot g \, d\mu \right|$$

$$+ \left| \int (\psi_{n_1}(f) - l(f)) \cdot \psi'(g) \, d\mu \right| < \varepsilon.$$

Hieraus folgt  $\int l(f) \cdot g \, d\mu = \int \psi(l(f)) \cdot g \, d\mu$  für alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  und  $g \in \mathcal{L}^1(\mathcal{A}, \mu)$  und somit

$l(f) = \psi(l(f))$   $\mu$ -fast überall.

Ist  $l(f) = f$   $\mu$ -fast überall, so liefert dies  $f = \psi(f)$   $\mu$ -fast überall. Ist  $\psi(f) = f$   $\mu$ -fast überall, so liefert dies auch  $\psi_n(f) = f$   $\mu$ -fast überall und daher auch  $l(f) = f$   $\mu$ -fast überall.

Sei  $f_n \in \mathcal{L}_+^\infty(\mathcal{A}, \mu)$  für alle  $n \in \mathbb{N}$ .  $(f_n)_{n \in \mathbb{N}}$  sei isoton mit

$\sup_{n \in \mathbb{N}} f_n \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  und  $l(f_n) = f_n$   $\mu$ -fast überall für alle  $n \in \mathbb{N}$ . Dann ist  $\psi(f_n) = f_n$   $\mu$ -fast überall für alle  $n \in \mathbb{N}$  und folglich  $\psi(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} f_n$   $\mu$ -fast überall und somit auch  $l(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} f_n$   $\mu$ -fast überall.

Nach (6.10)(a) ist für alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  und  $\vartheta \in \mathcal{O}$

$\int (f - \psi(f)) dP_\vartheta = 0$ , und deshalb ist auch

$\int (f - \psi_n(f)) dP_\vartheta = \int (f - \psi_n(f)) \frac{dP_\vartheta}{d\mu} d\mu = 0$  für alle  $n \in \mathbb{N}$ ,

was  $\int (f - l(f)) dP_\vartheta = 0$  nach sich zieht.

Sei  $l(f) = f$   $\mu$ -fast überall, dann ist  $l(f_+) \geq (l(f))_+ = f_+$

$\mu$ -fast überall. Man hat also  $\int (l(f_+) - f_+) dP_\vartheta = 0$  für

alle  $\vartheta \in \mathcal{O}$ , wobei der Integrand  $\mu$ -fast überall nicht negativ ist. Also ist  $l(f_+) = f_+$   $\mu$ -fast überall, was noch

$l(|f|) = |f|$   $\mu$ -fast überall liefert. Es gilt dann auch für

$f_1, f_2 \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  mit  $l(f_i) = f_i$ ,  $i = 1, 2$ ,

$l(\inf(f_1, f_2)) = \inf(f_1, f_2)$   $\mu$ -fast überall.

Es sei nun  $f \in \mathcal{L}_+^\infty(\mathcal{A}, \mu)$  mit  $l(f) = f$   $\mu$ -fast überall.  $\alpha > 0$

sei fest und  $f_n := \inf(n \cdot (f - \alpha)_+, 1)$  für alle  $n \in \mathbb{N}$ .  $(f_n)_{n \in \mathbb{N}}$

ist eine isotone Folge mit  $\sup_{n \in \mathbb{N}} f_n = 1_{\{f > \alpha\}}$  und  $l(f_n) = f_n$

$\mu$ -fast überall für alle  $n \in \mathbb{N}$ . Also ist  $l(1_{\{f > \alpha\}}) = 1_{\{f > \alpha\}}$

$\mu$ -fast überall und somit auch

$\psi(\{f > \alpha\}, \cdot) = \psi(1_{\{f > \alpha\}}) = 1_{\{f > \alpha\}}$   $\mu$ -fast überall.

$\{f > \alpha\}$  liegt daher in  $\mathcal{O}$ .

Es sind folglich alle  $f \in \mathcal{L}_+^\infty(\mathcal{A}, \mu)$  mit  $l(f) = f$   $\mu$ -fast

überall  $\mathcal{O}$ -meßbar. Dann sind aber auch alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  mit

$l(f) = f$   $\mu$ -fast überall  $\mathcal{O}$ -meßbar. Insbesondere sind für

alle  $f \in \mathcal{L}^\infty(\mathcal{A}, \mu)$  die Funktionen  $l(f)$   $\mathcal{O}$ -meßbar; denn

$l(f) = \psi(l(f))$   $\mu$ -fast überall liefert  $l(f) = \psi_n(l(f))$

$\mu$ -fast überall für alle  $n \in \mathbb{N}$  und dies  $l(f) = l(l(f))$

$\mu$ -fast überall.

Sei  $A \in \mathcal{A}$  und  $S \in \mathcal{S}$ , dann ist  $l(1_{A \cap S}) \leq l(1_S) = 1_S$   $\mu$ -fast überall und  $l(1_{A \cap (X \setminus S)}) \leq l(1_{X \setminus S}) = 1_{X \setminus S}$   $\mu$ -fast überall und weiter

$l(1_A) = l(1_A) \cdot 1_S + l(1_A) \cdot 1_{X \setminus S} = l(1_{A \cap S}) + l(1_{A \cap (X \setminus S)})$   $\mu$ -fast überall. Zusammen liefert das  $l(1_A) \cdot 1_S = l(1_{A \cap S})$   $\mu$ -fast überall. Für alle  $A \in \mathcal{A}$ ,  $S \in \mathcal{S}$  und  $\nu \in \mathcal{O}$  ist dann

$$\int_S 1_A dP_\nu = \int 1_{A \cap S} dP_\nu = \int l(1_{A \cap S}) dP_\nu = \int l(1_A) \cdot 1_S dP_\nu = \int_S l(1_A) dP_\nu.$$

Das bedeutet aber, daß ein bedingter  $(P_\nu)_{\nu \in \mathcal{O}}$ -Erwartungswert von  $1_A$  bezüglich  $\mathcal{S}$  existiert.

Dieser Satz entspricht Proposition 9 in [11]. Die Beweis-idee ist analog.

(6.12) Korollar.  $(P_\nu)_{\nu \in \mathcal{O}}$  sei durch ein  $\sigma$ -endliches Maß  $\mu$  dominiert. Dann ist  $\mathcal{I}$  eine suffiziente  $\sigma$ -Algebra für  $(P_\nu)_{\nu \in \mathcal{O}}$ .

Beweis. Es ist  $\mathcal{S} \subset \mathcal{I}$   $\mu$ -fast und  $\mathcal{S}$  suffizient für  $(P_\nu)_{\nu \in \mathcal{O}}$ . Da  $\mu$   $\sigma$ -endlich ist, ist dann auch  $\mathcal{I}$  suffizient für  $(P_\nu)_{\nu \in \mathcal{O}}$ .

(6.13) Korollar. Für ein beliebiges Experiment  $(X, \mathcal{A}, (P_\nu)_{\nu \in \mathcal{O}})$  ist ein Unterexperiment  $(X, \mathcal{I}, (P_\nu|_{\mathcal{I}})_{\nu \in \mathcal{O}})$  genau dann paarweise schwach suffizient, wenn es paarweise suffizient ist.

Beweis. Je zwei normierte Maße sind durch ein endliches Maß dominiert. (6.11) und (6.12) liefern dann die Behauptung.

## § 7. Schwache Suffizienz und paarweise schwache Suffizienz

Im folgenden seien  $(X, \alpha, (P_j)_{j \in \Theta})$  und  $(Y, \gamma, (Q_j)_{j \in \Theta})$  wieder zwei Experimente. Wir wollen zeigen, daß die paarweise schwache Suffizienz unter gewissen Voraussetzungen die schwache Suffizienz nach sich zieht. Daß auf diese Zusatzvoraussetzungen nicht ganz verzichtet werden darf, zeigt ein Beispiel in [14].

(7.1) Satz.  $(Y, \gamma, (Q_j)_{j \in \Theta})$  sei paarweise schwach suffizient für  $(X, \alpha, (P_j)_{j \in \Theta})$ , und  $(Q_j)_{j \in \Theta}$  sei schwach dominiert. Dann ist  $(Y, \gamma, (Q_j)_{j \in \Theta})$  schwach suffizient für  $(X, \alpha, (P_j)_{j \in \Theta})$ .

Beweis. Unter den Voraussetzungen des Satzes ist offensichtlich auch  $(X \times Y, \gamma', (P'_j | \gamma')_{j \in \Theta})$  schwach dominiert. Weiter ist  $\gamma'$  nach (6.7) und (6.8) paarweise suffizient für  $(P'_j)_{j \in \Theta}$ . Nach (3.5) ist  $\gamma'$  daher suffizient für  $(P'_j)_{j \in \Theta}$ . Wiederrum nach (6.7) und (6.8) ist dann  $(Y, \gamma, (Q_j)_{j \in \Theta})$  schwach suffizient für  $(X, \alpha, (P_j)_{j \in \Theta})$ .

Im Falle der Dominierttheit von  $(P_j)_{j \in \Theta}$  (und damit von  $(Q_j)_{j \in \Theta}$ ) ist dieser Satz in [14] bewiesen.

(7.2) Satz. Eines der Experimente  $(X, \alpha, (P_j)_{j \in \Theta})$  und  $(Y, \gamma, (Q_j)_{j \in \Theta})$  sei schwach dominiert. Sind die Experimente dann isomorph, so ist jedes für das andere schwach suffizient.

Beweis. Die Isomorphie zieht die paarweise schwache Suffizienz jedes Experimentes für das andere nach sich gemäß [14], Theorem 2.1. Insbesondere ist das schwach dominierte Experiment für das andere paarweise schwach suffizient.

(7.1) liefert, daß es dann auch schwach suffizient für das andere ist. Da schwach suffizient eine symmetrische Eigenschaft der beiden Experimente ist, folgt die Behauptung.

## § 8. f-Suffizienz

Csiszár [6] zeigt, daß sich paarweise Suffizienz auch informationstheoretisch charakterisieren läßt.

Im folgenden sei  $f$  eine konvexe Funktion auf  $]0, +\infty[$ .

Es wird vereinbart:

$$f(0) := \lim_{u \rightarrow 0} f(u), \quad 0 \cdot f\left(\frac{0}{0}\right) := 0,$$

$$0 \cdot f\left(\frac{a}{0}\right) := \lim_{\varepsilon \rightarrow 0} \varepsilon \cdot f\left(\frac{a}{\varepsilon}\right) = a \cdot \lim_{u \rightarrow \infty} \frac{f(u)}{u} \quad \text{für } 0 < a < +\infty.$$

$(X, \mathcal{A})$  sei ein Maßraum. Für zwei normierte Maße  $P_1, P_2$  auf  $\mathcal{A}$  mit Dichten  $p_1, p_2$  bezüglich eines  $\sigma$ -endlichen Maßes  $\lambda$  auf  $\mathcal{A}$  ist nach [6] das Integral

$$\int p_2(x) \cdot f\left(\frac{p_1(x)}{p_2(x)}\right) \lambda(dx) \quad \text{wohldefiniert (eventuell nicht}$$

endlich) und unabhängig von dem speziellen dominierenden

Maß  $\lambda$ . Dies führt zu der folgenden Definition.

(8.1) Definition. Für zwei normierte Maße  $P_1, P_2$  auf einem Maßraum  $(X, \mathcal{A})$ , die bezüglich eines  $\sigma$ -endlichen Maßes  $\lambda$  auf  $\mathcal{A}$  die Dichten  $p_1, p_2$  haben, heißt

$$I_f(P_1, P_2) := \int p_2(x) \cdot f\left(\frac{p_1(x)}{p_2(x)}\right) \lambda(dx)$$

$f$ -Divergenz von  $P_1$  nach  $P_2$ .

(8.2) Beispiel. Wir betrachten die Spezialfälle

$$f(u) = u \cdot \log(u) \quad \text{und} \quad f(u) = (u - 1)^2, \quad u \in ]0, +\infty[$$

Dann ist  $I_{u \cdot \log(u)}(P_1, P_2) = \int p_1(x) \cdot \log\left(\frac{p_1(x)}{p_2(x)}\right) \lambda(dx)$  und

$$I_{(u-1)^2}(P_1, P_2) = \int \frac{(p_1(x) - p_2(x))^2}{p_2(x)} \lambda(dx).$$

$(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  und  $(Y, \mathcal{L}, (Q_\vartheta)_{\vartheta \in \Theta})$  seien jetzt wieder zwei

Experimente. Je zwei normierte Maße sind stets durch ein

$\sigma$ -endliches Maß, zum Beispiel ihre Summe, dominiert.

(8.3) Definition.  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  heißt  $f$ -suffizient für  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ , wenn es einen schwachen Markovkern  $\varphi$  von  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  nach  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  gibt und für jede Wahl von  $\vartheta, \vartheta' \in \Theta$  gilt  $I_f(Q_\vartheta, Q_{\vartheta'}) < \infty$  und  $I_f(P_\vartheta, P_{\vartheta'}) = I_f(Q_\vartheta, Q_{\vartheta'})$ .

Csiszár zeigt im Anschluß an [6], Theorem (3.1), folgenden

(8.4) Satz. Für streng konvexes  $f$  ist ein Unterexperiment  $(X, \mathcal{A}', (P_\vartheta|\mathcal{A}')_{\vartheta \in \Theta})$  des Experimentes  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  genau dann  $f$ -suffizient für  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ , wenn  $\mathcal{A}'$  eine paarweise suffiziente Unter- $\sigma$ -Algebra von  $\mathcal{A}$  für  $(P_\vartheta)_{\vartheta \in \Theta}$  ist und für jede Wahl von  $\vartheta, \vartheta' \in \Theta$   $I_f(P_\vartheta|\mathcal{A}', P_{\vartheta'}|\mathcal{A}') < +\infty$  ist.

Hiervon ausgehend können wir eine Bedingung angeben, unter der  $f$ -Suffizienz und paarweise schwache Suffizienz äquivalent sind. Dies wird auch in [6] erwähnt. Wir wollen es hier ausführen. Zunächst benötigen wir ein Lemma.

(8.5) Lemma.  $\varphi$  sei ein schwacher Markovkern von dem Maßraum  $(X, \mathcal{A}, \mu)$  nach dem Maßraum  $(Y, \mathcal{Y})$ , und das Maß  $\mu'$  auf  $\mathcal{A} \otimes \mathcal{Y}$  sei wie in (6.5) definiert. Dann gilt für jede  $\mathcal{A}$ -meßbare Funktion  $g$ , für die  $\int g d\mu$  definiert ist,

$$\int_{A \times C} g \circ \text{pr}_1 d\mu' = \int_A \varphi(C, \cdot) \cdot g d\mu \quad \text{für alle } A \in \mathcal{A} \text{ und } C \in \mathcal{Y}.$$

Beweis. Da  $0 \leq 1_A \cdot \varphi(C, \cdot) \leq 1$   $\mu$ -fast überall für alle  $A \in \mathcal{A}$  und  $C \in \mathcal{Y}$  gilt, ist  $\int_A \varphi(C, \cdot) \cdot g d\mu$  für alle  $A \in \mathcal{A}$  und  $C \in \mathcal{Y}$  definiert.

Nun gilt für jede Indikatorfunktion  $1_B$ ,  $B \in \mathcal{A}$ ,

$$\int_A \varphi(C, \cdot) \cdot 1_B d\mu = \int_{(A \cap B) \times C} d\mu' = \int_{A \times C} 1_B \circ \text{pr}_1 d\mu'$$

für alle  $A \in \mathcal{A}$  und  $C \in \mathcal{Y}$ .

Nach der üblichen Schlußweise gilt dann auch für  $g$

$$\int_A \varphi(C, \cdot) \cdot g d\mu = \int_{A \times C} g \circ \text{pr}_1 d\mu' \quad \text{für alle } A \in \mathcal{A} \text{ und } C \in \mathcal{Y}.$$

Der Beweis geht auf Csizsár [6] zurück.

(8.6) Korollar.  $\varphi$  sei ein schwacher Markovkern von dem Maßraum  $(X, \mathcal{A}, \mu)$  nach dem Maßraum  $(Y, \mathcal{Y})$ .  $\nu$  sei auf  $\mathcal{A}$  ein endliches Maß, das eine Dichte bezüglich  $\mu$  hat.  $\mu'$  und  $\nu'$  seien auf  $\mathcal{A} \otimes \mathcal{Y}$  wie in (6.5) definiert.

Dann ist  $\frac{d\nu}{d\mu} \circ pr_1$  eine Dichte für  $\nu'$  bezüglich  $\mu'$ .

Beweis. (8.5) liefert für alle  $A \in \mathcal{A}$  und  $C \in \mathcal{Y}$

$$\int_{A \times C} \frac{d\nu}{d\mu} \circ pr_1 d\mu' = \int_A \varphi(C, \cdot) \cdot \frac{d\nu}{d\mu} d\mu = \int_A \varphi(C, \cdot) d\nu = \nu'(A \times C)$$

Da die Fortsetzung von endlichen Maßen von einem Semiring auf die davon erzeugte  $\sigma$ -Algebra eindeutig ist, folgt die Behauptung.

(8.7) Satz.  $f$  sei streng konvex.  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  ist genau dann  $f$ -suffizient für  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ , wenn  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  paarweise schwach suffizient für  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  ist und für jede Wahl von  $\vartheta, \vartheta' \in \Theta$   $I_f(Q_\vartheta, Q_{\vartheta'}) < +\infty$  ist.

Beweis. Wir betrachten wie in § 6 die  $\sigma$ -Algebra  $\mathcal{A} \otimes \mathcal{Y}$  und darauf die Familie von normierten Maßen  $(P'_\vartheta)_{\vartheta \in \Theta}$  und die Unter- $\sigma$ -Algebra  $\mathcal{Y}'$  von  $\mathcal{A} \otimes \mathcal{Y}$ . Nach (6.7) und (6.8) ist  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  genau dann paarweise schwach suffizient für  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ , wenn  $\mathcal{Y}'$  paarweise suffizient für  $(P'_\vartheta)_{\vartheta \in \Theta}$  ist. Nach (8.4) ist dies wiederum unter der Voraussetzung, daß für jede Wahl von  $\vartheta, \vartheta' \in \Theta$   $I_f(P'_\vartheta|_{\mathcal{Y}'}, P'_{\vartheta'}|_{\mathcal{Y}'}) < +\infty$  ist, äquivalent damit, daß  $(X \times Y, \mathcal{Y}', (P'_\vartheta|_{\mathcal{Y}'})_{\vartheta \in \Theta})$   $f$ -suffizient für  $(X \times Y, \mathcal{A} \otimes \mathcal{Y}, (P'_\vartheta)_{\vartheta \in \Theta})$  ist.

Wir haben also lediglich für alle  $\vartheta, \vartheta' \in \Theta$

$I_f(P'_\vartheta|_{\mathcal{Y}'}, P'_{\vartheta'}|_{\mathcal{Y}'}) = I_f(Q_\vartheta, Q_{\vartheta'})$  und  $I_f(P'_\vartheta, P'_{\vartheta'}) = I_f(P_\vartheta, P_{\vartheta'})$  nachzuweisen.

In der Tat: Seien  $\vartheta_1, \vartheta_2 \in \Theta$ ,  $\lambda := P_{\vartheta_1} + P_{\vartheta_2}$ ,  $p_1, p_2$  Dichten von  $P_{\vartheta_1}, P_{\vartheta_2}$  bezüglich  $\lambda$  und  $\varphi$  der nach Voraussetzung existierende schwache Markovkern von  $(\mathbf{X}, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  nach  $(Y, \mathcal{Y}, (Q_{\vartheta})_{\vartheta \in \Theta})$ .  $q_1, q_2$  seien Dichten von  $Q_{\vartheta_1}, Q_{\vartheta_2}$  bezüglich  $\mu := Q_{\vartheta_1} + Q_{\vartheta_2}$ , und  $\lambda'$  sei wie in (6.5) definiert. Es ist  $\mu = \lambda' \circ \text{pr}_2^{-1}$ . Wegen

$$\begin{aligned} P'_{\vartheta_1}(\mathbf{X} \times \mathbf{C}) &= Q_{\vartheta_1}(\mathbf{C}) = \int_{\mathbf{C}} q_1 \, dQ_{\vartheta_1} + \int_{\mathbf{C}} q_1 \, dQ_{\vartheta_2} \\ &= \int_{\mathbf{C}} q_1 \, dP'_{\vartheta_1} \circ \text{pr}_2^{-1} + \int_{\mathbf{C}} q_1 \, dP'_{\vartheta_2} \circ \text{pr}_2^{-1} \\ &= \int_{\mathbf{X} \times \mathbf{C}} q_1 \circ \text{pr}_2 \, dP'_{\vartheta_1} + \int_{\mathbf{X} \times \mathbf{C}} q_1 \circ \text{pr}_2 \, dP'_{\vartheta_2} = \int_{\mathbf{X} \times \mathbf{C}} q_1 \circ \text{pr}_2 \, d\lambda' \\ &= \int_{\mathbf{X} \times \mathbf{C}} q_1 \circ \text{pr}_2 \, d\lambda' | \mathcal{Y}' \text{ für alle } \mathbf{C} \in \mathcal{Y}' \text{ und } i = 1, 2 \end{aligned}$$

sind  $q_1 \circ \text{pr}_2, q_2 \circ \text{pr}_2$  Dichten von  $P'_{\vartheta_1} | \mathcal{Y}', P'_{\vartheta_2} | \mathcal{Y}'$  bezüglich  $\lambda' | \mathcal{Y}'$ .

Also ist

$$\begin{aligned} I_f(P'_{\vartheta_1} | \mathcal{Y}', P'_{\vartheta_2} | \mathcal{Y}') &= \int q_2 \circ \text{pr}_2 \cdot f\left(\frac{q_1(\cdot)}{q_2(\cdot)}\right) \circ \text{pr}_2 \, d\lambda' | \mathcal{Y}' \\ &= \int q_2 \cdot f\left(\frac{q_1(\cdot)}{q_2(\cdot)}\right) \, d\lambda' \circ \text{pr}_2^{-1} = \int q_2 \cdot f\left(\frac{q_1(\cdot)}{q_2(\cdot)}\right) \, d\mu = I_f(Q_{\vartheta_1}, Q_{\vartheta_2}). \end{aligned}$$

Gemäß (8.6) ist

$$\begin{aligned} I_f(P'_{\vartheta_1}, P'_{\vartheta_2}) &= \int p_2 \circ \text{pr}_1 \cdot f\left(\frac{p_1(\cdot)}{p_2(\cdot)}\right) \circ \text{pr}_1 \, d\lambda' \\ &= \int p_2 \cdot f\left(\frac{p_1(\cdot)}{p_2(\cdot)}\right) \, d\lambda' \circ \text{pr}_1^{-1} = \int p_2 \cdot f\left(\frac{p_1(\cdot)}{p_2(\cdot)}\right) \, d\lambda = I_f(P_{\vartheta_1}, P_{\vartheta_2}). \end{aligned}$$

(8.8) Satz.  $f$  sei streng konvex.  $(Y, \mathcal{Y}, (Q_{\vartheta})_{\vartheta \in \Theta})$  sei schwach dominiert, und für alle  $\vartheta, \vartheta' \in \Theta$  sei  $I_f(Q_{\vartheta}, Q_{\vartheta'}) < +\infty$ .

Dann ist  $(Y, \mathcal{Y}, (Q_{\vartheta})_{\vartheta \in \Theta})$  genau dann schwach suffizient für  $(\mathbf{X}, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$ , wenn  $(Y, \mathcal{Y}, (Q_{\vartheta})_{\vartheta \in \Theta})$   $f$ -suffizient für  $(\mathbf{X}, \mathcal{A}, (P_{\vartheta})_{\vartheta \in \Theta})$  ist.

Beweis. (8.7) ergibt zusammen mit (7.1) die Behauptung.

## § 9. Ein Isomorphiesatz

(9.1) Definition. Zwei Experimente  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  und  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  heißen maisomorph, wenn es eine bijektive Abbildung  $T: \mathcal{A}((P_\vartheta)_{\vartheta \in \Theta}) \rightarrow \mathcal{Y}((Q_\vartheta)_{\vartheta \in \Theta})$  gibt mit folgenden Eigenschaften:

- (1)  $T(\widetilde{X \setminus A}) = \widetilde{Y \setminus T(A)}$  fr alle  $A \in \mathcal{A}$ ,
- (2)  $T(\bigcap_{n=1}^{\infty} A_n) = \bigcap_{n=1}^{\infty} T(A_n)$  fr jede Folge  $(A_n)_{n \in \mathbb{N}}$  von Mengen aus  $\mathcal{A}$ ,
- (3)  $Q_\vartheta(T(A)) = P_\vartheta(A)$  fr alle  $A \in \mathcal{A}$  und  $\vartheta \in \Theta$ .

Wenn zwei Experimente schwach dominiert sind, haben sie nach (5.3) minimalsuffiziente Unterexperimente. Wir wollen zeigen, da diese maisomorph sind, wenn das eine Ausgangs-experiment fr das andere schwach suffizient ist.

$(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  und  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  seien zwei Experimente.  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  sei schwach suffizient fr  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ . Das Experiment  $(X \times Y, \mathcal{A} \otimes \mathcal{Y}, (P'_\vartheta)_{\vartheta \in \Theta})$  sei wie in § 6 gebildet. Wir bentigen zunchst das folgende Lemma:

(9.2) Lemma.  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$  und  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  seien schwach dominiert, und  $(Y, \mathcal{Y}, (Q_\vartheta)_{\vartheta \in \Theta})$  sei schwach suffizient fr  $(X, \mathcal{A}, (P_\vartheta)_{\vartheta \in \Theta})$ . Dann ist  $(X \times Y, \mathcal{A} \otimes \mathcal{Y}, (P'_\vartheta)_{\vartheta \in \Theta})$  schwach dominiert.

Beweis. Nach Voraussetzung und (6.6), (6.7) und (6.8) sind die Unter- $\sigma$ -Algebren  $\mathcal{A}'$  und  $\mathcal{Y}'$  von  $\mathcal{A} \otimes \mathcal{Y}$  suffizient fr  $(P'_\vartheta)_{\vartheta \in \Theta}$ . Fr jedes  $D \in \mathcal{A} \otimes \mathcal{Y}$  gibt es also von  $1_D$  einen bedingten  $(P'_\vartheta)_{\vartheta \in \Theta}$ -Erwartungswert  $E(1_D | \mathcal{A}')$  bezglich  $\mathcal{A}'$  und einen  $E(1_D | \mathcal{Y}')$  bezglich  $\mathcal{Y}'$ .

$\mu$  sei nun ein zu  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \Theta}$  äquivalentes schwach dominierendes Maß für  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \Theta}$ . Nach (2.3) gibt es eine Familie von Mengen aus  $\mathcal{O}$   $(X_i)_{i \in I}$ ,  $I \neq \emptyset$ , so, daß jedes  $X_i$ ,  $i \in I$ , positives  $\sigma$ -endliches  $\mu$ -Maß hat,  $\mu(X_i \cap X_{i'}) = 0$  ist für alle  $i, i' \in I$  mit  $i \neq i'$  und aus  $\mu(A \cap X_i) = 0$ , wobei  $A \in \mathcal{O}$  ist, für alle  $i \in I$  folgt  $\mu(A) = 0$ .

Da  $\mu'(X_i \times Y) = \mu(X_i)$  ist für alle  $i \in I$ , hat jedes  $X_i \times Y$ ,  $i \in I$ , positives  $\sigma$ -endliches  $\mu'$ -Maß, und ferner ist  $P_{\mathfrak{J}}((X_i \times Y) \cap (X_{i'} \times Y)) = 0$  für alle  $i, i' \in I$  mit  $i \neq i'$  und  $\mathfrak{J} \in \Theta$ . Sei weiter  $D \in \mathcal{O} \otimes \mathcal{Y}$  und  $P_{\mathfrak{J}}(D \cap (X_i \times Y)) = 0$  für alle  $i \in I$  und  $\mathfrak{J} \in \Theta$ , so ergibt sich für alle  $\mathfrak{J} \in \Theta$  und  $i \in I$   $P_{\mathfrak{J}}(\{E(1_D | \mathcal{O}') > 0\} \cap (X_i \times Y)) = 0$ . Da  $\{E(1_D | \mathcal{O}') > 0\} = A \times Y$  ist für ein geeignetes  $A \in \mathcal{O}$ , folgt wegen der Eigenschaften von  $(X_i)_{i \in I}$   $P_{\mathfrak{J}}(\{E(1_D | \mathcal{O}') > 0\}) = 0$  für alle  $\mathfrak{J} \in \Theta$ , was  $P_{\mathfrak{J}}(D) = 0$  für alle  $\mathfrak{J} \in \Theta$  nach sich zieht.

Nach (8.6) ist  $(P_{\mathfrak{J}})_{\mathfrak{J} \in \Theta}$  bezüglich  $\mu'$  absolut stetig. Da  $\mu' \cdot X_i \times Y$ ,  $i \in I$ , ein  $\sigma$ -endliches Maß ist, ist also  $(P_{\mathfrak{J}}^{X_i \times Y})_{\mathfrak{J} \in \Theta}$  durch das  $\sigma$ -endliche Maß  $\mu' \cdot X_i \times Y$  dominiert,  $i \in I$ . Für alle  $i \in I$  gibt es daher eine abzählbare Teilmenge  $\Theta_i \subset \Theta$  so, daß  $(P_{\mathfrak{J}}^{X_i \times Y})_{\mathfrak{J} \in \Theta_i}$  äquivalent zu  $(P_{\mathfrak{J}}^{X_i \times Y})_{\mathfrak{J} \in \Theta}$  ist. Bei geeigneter Wahl von Elementen  $c_{i,\mathfrak{J}} \in \mathbb{R}_+$ ,  $i \in I$  und  $\mathfrak{J} \in \Theta_i$ , ist für alle  $i \in I$   $\nu_i := \sum_{\mathfrak{J} \in \Theta_i} c_{i,\mathfrak{J}} P_{\mathfrak{J}}^{X_i \times Y}$  ein zu  $(P_{\mathfrak{J}}^{X_i \times Y})_{\mathfrak{J} \in \Theta}$  äquivalentes normiertes Maß. Es sei

$$\nu(D) := \begin{cases} \sum_{i \in J} \nu_i(D \cap (X_i \times Y)), & J := \{i \in I : \nu_i(D \cap (X_i \times Y)) > 0\}, \\ & \text{wenn } J \text{ abzählbar und nicht leer} \\ & \text{ist und die Summe konvergiert,} \\ 0, & \text{wenn } J \text{ leer ist,} \\ \infty & \text{sonst} \end{cases}$$

für alle  $D \in \mathcal{O} \otimes \mathcal{Y}$ .

Wegen der Eigenschaften der Familie  $(X_i \times Y)_{i \in I}$  ist  $\nu$  ein zu  $(P'_j)_{j \in \mathcal{C}}$  äquivalentes Maß mit der endlichen Teilmengeneigenschaft, und da für jedes  $j \in \mathcal{C}$   $P'_j(X_i \times Y) = P_j(X_i) > 0$  nur für höchstens abzählbar viele  $i \in I$  gelten kann, hat jedes  $P'_j$ ,  $j \in \mathcal{C}$ , eine Dichte bezüglich  $\nu$  nach dem Satz von Radon-Nikodym.

Die Behauptung des Lemmas folgt, wenn wir zeigen, daß  $\nu$  lokalisierbar ist. Nach [15], Lemma S. 262, genügt dazu der Nachweis, daß für jede Familie  $(D_i)_{i \in I}$  von Mengen aus  $\mathcal{A} \otimes \mathcal{X}$  mit  $D_i \subset X_i \times Y$  für alle  $i \in I$   $\sup_{i \in I} \widetilde{D}_i^\nu$  existiert.

Aus  $D_i \subset X_i \times Y$ ,  $i \in I$ , folgt

$\nu(D_i) = \nu_i(D_i) = \sum_{j \in \mathcal{C}_i} c_{i,j} P'_j X_i \times Y(D_i)$ . Zu jedem  $P'_j X_i \times Y$ ,  $i \in I$  und  $j \in \mathcal{C}_i$ , gibt es Mengen  $A_{k,1}^{i,j} \in \mathcal{A}$  und  $C_{k,1}^{i,j} \in \mathcal{X}$ ,  $k, 1 \in \mathbb{N}$ , derart, daß  $D_{i,j} := \bigcap_{k=1}^{\infty} \left( \bigcup_{l=1}^{\infty} (A_{k,l}^{i,j} \times C_{k,l}^{i,j}) \right)$  die Eigenschaften  $D_{i,j} \supset D_i$  für alle  $i \in I$  und  $j \in \mathcal{C}_i$  und

$P'_j(D_{i,j}) = P'_j(D_i)$  für alle  $i \in I$  und  $j \in \mathcal{C}_i$  besitzt. Es sei o.E.d.A.  $D_{i,j} \subset X_i \times Y$  für alle  $j \in \mathcal{C}_i$ ,  $i \in I$ .

Sei  $D_i^a := \bigcap_{j \in \mathcal{C}_i} D_{i,j}$  für alle  $i \in I$ . Dann ist  $D_i^a \supset D_i$  und  $P'_j X_i \times Y(D_i^a) = P'_j X_i \times Y(D_i)$  für alle  $i \in I$  und  $j \in \mathcal{C}_i$ . Also ist  $\nu(D_i^a) = \nu(D_i)$  für alle  $i \in I$ . Es bleibt also zu zeigen, daß  $\sup_{i \in I} \widetilde{D}_i^\nu$  existiert, wenn jedes  $D_i$ ,  $i \in I$ , vom Typ  $D_i^a$  ist. Dies führen wir in vier Schritten aus.  $J$  sei eine nicht leere Indexmenge.

1. Für alle  $j \in J$  sei  $A_j \in \mathcal{A}$ . Dann existiert  $\sup_{j \in J} \widetilde{A_j \times Y}^\nu$ .

Denn sei  $A \in \sup_{j \in J} \widetilde{A_j}^\mu$ , dann ist  $A_j \subset A$   $\mu$ -fast für alle  $j \in J$  und folglich auch  $(P'_j)_{j \in \mathcal{C}}$ -fast. Daher ist  $A_j \times Y \subset A \times Y$   $(P'_j)_{j \in \mathcal{C}}$ -fast und folglich auch  $\nu$ -fast für alle  $j \in J$ .

Angenommen es existiert ein  $D \in \mathcal{A} \otimes \mathcal{X}$  mit  $A_j \times Y \subset D \subsetneq A \times Y$

$\nu$ -fast für alle  $j \in J$ . Dann ist  $1_{A_j \times Y} \in E(1_D | \mathcal{A}') \subset 1_{A \times Y}$  auf einer Menge mit von Null verschiedenem  $\nu$ -Maß für alle  $j \in J$  und daher  $A_j \times Y \subset \{E(1_D | \mathcal{A}') = 1\} \subsetneq A \times Y$   $\nu$ -fast für alle  $j \in J$ . Nun ist aber  $\{E(1_D | \mathcal{A}') = 1\} = A' \times Y$  für ein geeignetes  $A' \in \mathcal{A}$ . Unsere Annahme steht daher im Widerspruch zur Minimalität von  $A$ . Also ist  $\widetilde{A \times Y}^\nu = \sup_{j \in J} \widetilde{A_j \times Y}^\nu$ .

2. Für alle  $j \in J$  sei  $C_j \in \mathcal{X}$ . Dann existiert  $\sup_{j \in J} \widetilde{X \times C_j}^\nu$ .

Dies beweist man analog zu 1.

3. Für alle  $j \in J$  sei  $A_j \in \mathcal{A}$  und  $C_j \in \mathcal{X}$ . Dann existiert  $\sup_{j \in J} \widetilde{A_j \times C_j}^\nu$ .

Denn nach 1. und 2. gibt es Mengen  $D, E \in \mathcal{A} \otimes \mathcal{X}$  mit

$$\widetilde{D}^\nu = \sup_{j \in J} \widetilde{A_j \times Y}^\nu \text{ und } \widetilde{E}^\nu = \sup_{j \in J} \widetilde{X \times C_j}^\nu. \text{ Es ist dann } \widetilde{D \cap E}^\nu = \sup_{j \in J} \widetilde{A_j \times C_j}^\nu.$$

4. Für alle  $j \in J$  und  $n \in \mathbb{N}$  sei  $D_{j,n} \in \mathcal{A} \otimes \mathcal{X}$ , und für alle  $n \in \mathbb{N}$  existiere  $\sup_{j \in J} \widetilde{D_{j,n}}^\nu$ . Dann existieren

$$\sup_{j \in J} \bigcirc_{n=1}^\infty \widetilde{D_{j,n}}^\nu \text{ und } \sup_{j \in J} \bigcirc_{n=1}^\infty \widetilde{D_{j,n}}^\nu.$$

Sei nämlich  $D_n \in \sup_{j \in J} \widetilde{D_{j,n}}^\nu$ ,  $n \in \mathbb{N}$ , dann ist

$$\bigcirc_{n=1}^\infty \widetilde{D_n}^\nu = \sup_{j \in J} \bigcirc_{n=1}^\infty \widetilde{D_{j,n}}^\nu \text{ und } \bigcirc_{n=1}^\infty \widetilde{D_n}^\nu = \sup_{j \in J} \bigcirc_{n=1}^\infty \widetilde{D_{j,n}}^\nu.$$

Mit 1. bis 4. sind die Mengen vom Typ  $D_1^a$  erfaßt. Damit folgt die Lokalisierbarkeit von  $\nu$ .

(9.3) Satz.  $(Y, \mathcal{X}, (Q_\beta)_{\beta \in \Theta})$  sei schwach suffizient für  $(X, \mathcal{A}, (P_\beta)_{\beta \in \Theta})$ . Beide Experimente seien schwach dominiert. Dann sind die (nach (5.3) existierenden) minimalsuffizienten Unterexperimente dieser Experimente maßisomorph.

Beweis.  $\mu$  sei ein zu  $(P_\beta)_{\beta \in \Theta}$  äquivalentes schwach dominierendes Maß für  $(P_\beta)_{\beta \in \Theta}$ . Nach (9.2) gibt es ein zu  $(P'_\beta)_{\beta \in \Theta}$

äquivalentes schwach dominierendes Maß  $\nu$  für  $(P'_j)_{j \in \Theta}$ . Nach (5.3) gibt es daher eine für  $(P'_j)_{j \in \Theta}$  minimalsuffiziente Unter- $\sigma$ -Algebra  $\mathcal{D}$  von  $\mathcal{O} \otimes \mathcal{I}$ .  $\mathcal{O}^*$  bzw.  $\mathcal{I}^*$  seien für  $(P_j)_{j \in \Theta}$  bzw.  $(Q_j)_{j \in \Theta}$  minimalsuffiziente Unter- $\sigma$ -Algebren von  $\mathcal{O}$  bzw.  $\mathcal{I}$ .

Da  $\mathcal{O}^*$  und  $\mathcal{I}^*$  suffizient für  $(P'_j)_{j \in \Theta}$  sind nach (6.6), (6.7) und (6.8), sind auch die Unter- $\sigma$ -Algebren

$\mathcal{O}^{**} := \{A \times Y : A \in \mathcal{O}^*\}$  und  $\mathcal{I}^{**} := \{X \times C : C \in \mathcal{I}^*\}$  von  $\mathcal{O} \otimes \mathcal{I}$  suffizient für  $(P'_j)_{j \in \Theta}$ . Daher gilt  $\mathcal{D} \subset \mathcal{O}^{**}$  und  $\mathcal{D} \subset \mathcal{I}^{**}$   $\nu$ -fast.

$\mathcal{D}^* := \{A \in \mathcal{O}^* : \text{Es gibt ein } S \in \mathcal{D} \text{ mit } \nu(S \Delta (A \times Y)) = 0\}$  ist eine Unter- $\sigma$ -Algebra von  $\mathcal{O}^*$  mit  $\mathcal{D}^{**} := \{S \times Y : S \in \mathcal{D}^*\} \subset \mathcal{D}$   $\nu$ -fast. Da  $\mathcal{D} \subset \mathcal{O}^{**}$   $\nu$ -fast ist, gibt es zu jeder Menge  $S \in \mathcal{D}$  ein  $A_S \in \mathcal{O}^*$  mit  $\nu(S \Delta (A_S \times Y)) = 0$ . Also ist  $A_S \in \mathcal{D}^*$ .

Hieraus ergibt sich dann, daß es zu jeder  $\mathcal{D}$ -meßbaren Funktion  $f$  eine  $\mathcal{D}^*$ -meßbare Funktion  $f^*$  gibt mit  $f = f^* \circ \text{pr}_1$   $\nu$ -fast überall. Für  $A \in \mathcal{O}$  sei  $E(1_{A \times Y} | \mathcal{D})$  ein bedingter  $(P'_j)_{j \in \Theta}$ -Erwartungswert von  $1_{A \times Y}$  bezüglich  $\mathcal{D}$ . Dann gilt

$$\begin{aligned} \int_{S^*} 1_A dP_{\mathcal{D}} &= \int_{S^* \times Y} 1_{A \times Y} dP_{\mathcal{D}} = \int_{S^* \times Y} E(1_{A \times Y} | \mathcal{D}) dP_{\mathcal{D}} \\ &= \int_{S^* \times Y} (E(1_{A \times Y} | \mathcal{D}))^* \circ \text{pr}_1 dP'_j = \int_{S^*} (E(1_{A \times Y} | \mathcal{D}))^* dP'_j \circ \text{pr}_1^{-1} \\ &= \int_{S^*} (E(1_{A \times Y} | \mathcal{D}))^* dP_{\mathcal{D}} \text{ für alle } S^* \in \mathcal{D}^* \text{ und } j \in \Theta. \end{aligned}$$

Daraus folgt, daß  $\mathcal{D}^*$  suffizient für  $(P_j)_{j \in \Theta}$  ist. Daher ist  $\mathcal{D}^* = \mathcal{O}^* \mu$ -fast, und das zieht  $\mathcal{D}^{**} = \mathcal{D} = \mathcal{O}^{**}$   $\nu$ -fast nach sich. In analoger Weise erhält man  $\mathcal{D} = \mathcal{I}^{**}$   $\nu$ -fast.

Zusammen hat man  $\mathcal{O}^{**} = \mathcal{I}^{**}$   $\nu$ -fast, d.h. zu jedem  $A \in \mathcal{O}^*$  gibt es ein  $C \in \mathcal{I}^*$  mit  $\nu((A \times Y) \Delta (X \times C)) = 0$  und zu jedem  $C \in \mathcal{I}^*$  gibt es ein  $A \in \mathcal{O}^*$  mit  $\nu((A \times Y) \Delta (X \times C)) = 0$ .

$A \in \mathcal{A}^*$  seien auf diese Weise  $C_1, C_2 \in \mathcal{X}^*$  zugeordnet.

Dann ist

$$\begin{aligned} Q_{\mathcal{Y}}(C_1 \triangle C_2) &= P'_{\mathcal{Y}}(\mathcal{X} \times (C_1 \triangle C_2)) \\ &= P'_{\mathcal{Y}}(((\mathcal{X} \times C_1) \triangle (A \times Y)) \triangle ((A \times Y) \triangle (\mathcal{X} \times C_2))) \\ &\leq P'_{\mathcal{Y}}((\mathcal{X} \times C_1) \triangle (A \times Y)) + P'_{\mathcal{Y}}((\mathcal{X} \times C_2) \triangle (A \times Y)) = 0 \end{aligned}$$

für alle  $\mathcal{Y} \in \mathcal{O}$ . Ebenso ist  $P_{\mathcal{Y}}(A_1 \triangle A_2) = 0$ , wenn  $A_1$  und  $A_2$  dasselbe  $C \in \mathcal{X}^*$  zugeordnet ist.

Man hat also eine bijektive Abbildung  $T$  von  $\mathcal{A}^*((P_{\mathcal{Y}})_{\mathcal{Y} \in \mathcal{O}})$  auf  $\mathcal{X}^*((Q_{\mathcal{Y}})_{\mathcal{Y} \in \mathcal{O}})$ .

Aus  $\nu((A \times Y) \triangle (\mathcal{X} \times C)) = 0$  folgt

$\nu(((\mathcal{X} \setminus A) \times Y) \triangle (\mathcal{X} \times (Y \setminus C))) = 0$  für  $A \in \mathcal{A}^*$  und  $C \in \mathcal{X}^*$ . Damit hat man  $T(\widetilde{\mathcal{X} \setminus A}) = \widetilde{Y} \setminus T(\widetilde{A})$ .

Aus  $\nu((A_n \times Y) \triangle (\mathcal{X} \times C_n)) = 0$  für alle  $A_n \in \mathcal{A}^*$  und  $C_n \in \mathcal{X}^*$ ,  $n \in \mathbb{N}$ , folgt  $\nu(((\bigcup_{n=1}^{\infty} A_n) \times Y) \triangle (\mathcal{X} \times (\bigcup_{n=1}^{\infty} C_n))) = 0$ , und daher ist  $T(\bigcup_{n=1}^{\infty} \widetilde{A_n}) = \bigcup_{n=1}^{\infty} T(\widetilde{A_n})$ .

Schließlich gilt für  $A \in \mathcal{A}^*$  und  $C \in T(\widetilde{A})$

$P_{\mathcal{Y}}(A) = P'_{\mathcal{Y}}(A \times Y) = P'_{\mathcal{Y}}(A \times C)$ , da  $P'_{\mathcal{Y}}(A \times (Y \setminus C)) = 0$  ist, und  $Q_{\mathcal{Y}}(C) = P'_{\mathcal{Y}}(\mathcal{X} \times C) = P'_{\mathcal{Y}}(A \times C)$ , da  $P'_{\mathcal{Y}}((\mathcal{X} \setminus A) \times C) = 0$  ist, für alle  $\mathcal{Y} \in \mathcal{O}$ .

$(\mathcal{X}, \mathcal{A}^*, (P_{\mathcal{Y}}|_{\mathcal{A}^*})_{\mathcal{Y} \in \mathcal{O}})$  und  $(Y, \mathcal{X}^*, (Q_{\mathcal{Y}}|_{\mathcal{X}^*})_{\mathcal{Y} \in \mathcal{O}})$  sind also maßisomorph.

Die Präposition 12 in [11] entspricht dem eben bewiesenen Satz, dort wird allerdings topologisch argumentiert.

### III. Entscheidungstheoretische Charakterisierung der schwachen Suffizienz

Im letzten Teil dieser Arbeit sollen die erhaltenen Ergebnisse mit bereits bekannten Ergebnissen von Le Cam [11] und Heyer [8] verglichen werden. Zu diesem Zweck benötigen wir einige Hilfsmittel, die wir im folgenden Paragraphen bereitstellen werden.

#### § 10. Hilfssätze über spezielle lineare und bilineare Abbildungen

$\mathcal{F}$  und  $\mathcal{X}$  seien Vektorverbände reeller Funktionen mit 1 auf den (nicht leeren) Mengen  $X$  bzw.  $Y$ .  $\mathcal{A}(\mathcal{F})$  bzw.  $\mathcal{A}(\mathcal{X})$  seien die kleinsten  $\sigma$ -Algebren über  $X$  bzw.  $Y$  bezüglich derer alle Funktionen aus  $\mathcal{F}$  bzw. alle Funktionen aus  $\mathcal{X}$  meßbar sind.

(10.1) Lemma.  $T$  sei eine positive, normierte lineare Abbildung  $T: \mathcal{F} \rightarrow \mathcal{X}$  mit  $T(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} T(f_n)$  für jede isotone Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen aus  $\mathcal{F}_+$  mit  $\sup_{n \in \mathbb{N}} f_n \in \mathcal{F}$ . Dann gibt es genau einen Markovkern  $\mu$  von  $(Y, \mathcal{A}(\mathcal{X}))$  nach  $(X, \mathcal{A}(\mathcal{F}))$  mit

$$T(f)(\bar{y}) = \int f(x) \mu(dx, y) \quad \text{für alle } f \in \mathcal{F} \text{ und } y \in Y.$$

Beweis. Für jedes  $y \in Y$  sei  $\bar{y}$  der Auswertungshomomorphismus  $\bar{y}: \mathcal{X} \ni h \mapsto h(y) \in \mathbb{R}$ .  $\bar{y}$  ist positiv, normiert und linear. Damit ist  $\bar{y} \circ T$  eine positive, normierte Linearform auf  $\mathcal{F}$ .

Sei  $(f_n)_{n \in \mathbb{N}}$  eine isotone Folge von Funktionen aus  $\mathcal{F}_+$  mit  $\sup_{n \in \mathbb{N}} f_n \in \mathcal{F}$ , so gilt nach der Voraussetzung  $T(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} T(f_n)$ . Diese Gleichung ist äquivalent mit der Gleichung  $\bar{y} \circ T(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} \bar{y} \circ T(f_n)$  für alle  $y \in Y$ .

Nach dem Satz von Daniell-Stone ([21, Satz 39.4]) gibt es

daher für jedes  $y \in Y$  genau ein normiertes Maß  $\mu(\cdot, y)$  auf  $\mathcal{A}(\mathcal{F})$  mit  $\mathcal{F} \subset \mathcal{L}^1(\mathcal{A}(\mathcal{F}), \mu(\cdot, y))$  und

$$T(f)(y) = \bar{y} \cdot T(f) = \int f(x) \mu(dx, y) \text{ für alle } f \in \mathcal{F}.$$

Es ist noch zu zeigen, daß  $\mu(A, \cdot)$   $\mathcal{A}(\mathcal{X})$ -meßbar ist für jedes  $A \in \mathcal{A}(\mathcal{F})$ . Sei

$$\mathcal{O}(\mathcal{F}) := \{Q \in \mathcal{P}(\mathcal{X}) : 1_Q = \sup_{n \in \mathbb{N}} f_n, (f_n)_{n \in \mathbb{N}} \text{ isoton, } f_n \in \mathcal{F}_+\}.$$

Nach [2], Lemma 39.3, ist  $\mathcal{O}(\mathcal{F})$  durchschnittsstabil und erzeugt  $\mathcal{A}(\mathcal{F})$ .

Für jedes  $G \in \mathcal{O}(\mathcal{F})$  gibt es eine isotone Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen aus  $\mathcal{F}_+$  mit  $1_G = \sup_{n \in \mathbb{N}} f_n$ , und damit ist

$$\int 1_G(x) \mu(dx, y) = \int \sup_{n \in \mathbb{N}} f_n(x) \mu(dx, y) = \sup_{n \in \mathbb{N}} \int f_n(x) \mu(dx, y)$$

für alle  $y \in Y$ . Nun ist aber  $\int f_n(x) \mu(dx, \cdot) \in \mathcal{X}$  für alle  $n \in \mathbb{N}$ , und daher ist  $\mu(G, \cdot)$   $\mathcal{A}(\mathcal{X})$ -meßbar.

Sei  $\mathcal{O} := \{A \in \mathcal{A}(\mathcal{F}) : \mu(A, \cdot) \text{ ist } \mathcal{A}(\mathcal{X})\text{-meßbar}\}$ . Für  $A_1, A_2 \in \mathcal{O}$  mit  $A_1 \supset A_2$  ist  $\mu(A_1 \setminus A_2, \cdot) = \mu(A_1, \cdot) - \mu(A_2, \cdot)$   $\mathcal{A}(\mathcal{X})$ -meßbar.

Für jede Folge paarweiser disjunkter Mengen aus  $\mathcal{O}$   $(A_n)_{n \in \mathbb{N}}$  ist  $\mu(\bigcup_{n=1}^{\infty} A_n, \cdot) = \sum_{n=1}^{\infty} \mu(A_n, \cdot)$   $\mathcal{A}(\mathcal{X})$ -meßbar. Da außerdem wegen  $1 \in \mathcal{F}$   $\mathcal{X} \in \mathcal{O}$  ist, ist  $\mathcal{O}$  ein Dynkin-System. Weil  $\mathcal{O}(\mathcal{F})$  durchschnittsstabil ist und  $\mathcal{O}(\mathcal{F}) \subset \mathcal{O} \subset \mathcal{A}(\mathcal{F})$  gilt, ist  $\mathcal{O} = \mathcal{A}(\mathcal{F})$

([2], Satz 2.3). Also ist  $\mu(A, \cdot)$   $\mathcal{A}(\mathcal{X})$ -meßbar für alle  $A \in \mathcal{A}(\mathcal{F})$ .

(10.2) Korollar.  $X$  sei ein kompakter topologischer Raum und

$T$  eine positive, normierte lineare Abbildung  $T: \mathcal{C}(X) \rightarrow X$ .

Dann gibt es genau einen Markovkern  $\mu$  von  $(Y, \mathcal{A}(X))$  nach

$(X, \mathcal{A}(\mathcal{C}(X)))$  mit  $T(f)(y) = \int f(x) \mu(dx, y)$  für alle  $f \in \mathcal{C}(X)$

und  $y \in Y$ .

Beweis. Aus [2], Korollar 43.2, folgt, daß jedes  $T$  der

obigen Gestalt die in (10.1) geforderte Stetigkeitsbedingung erfüllt.

Im folgenden sei  $(Y, \mathfrak{X}, \nu)$  ein Maßraum, wobei  $\nu$  lokalisierbar sei und die endliche Teilmengeneigenschaft habe, und  $\mathfrak{F}$  sei ein Vektorverband beschränkter reeller Funktionen mit 1. Eine positive Bilinearform  $M$  auf  $\mathfrak{F} \times L^1(\mathfrak{X}, \nu)$  heißt normiert, wenn für alle  $g \in \mathcal{L}_+^1(\mathfrak{X}, \nu)$  gilt  $M(1, \tilde{g}) = \|\tilde{g}\|_1$ .

(10.3) Lemma. Die Gesamtheit der positiven, normierten Bilinearformen auf  $\mathfrak{F} \times L^1(\mathfrak{X}, \nu)$  ist isomorph zu der Gesamtheit der positiven, normierten linearen Abbildungen von  $\mathfrak{F}$  in  $L^\infty(\mathfrak{X}, \nu)$ .

Beweis.  $M$  sei eine positive, normierte Bilinearform auf  $\mathfrak{F} \times L^1(\mathfrak{X}, \nu)$ . Sei  $f_0 \in \mathfrak{F}$ , dann ist  $M(f_0, \cdot)$  eine Linearform auf  $L^1(\mathfrak{X}, \nu)$ . Nach den Voraussetzungen über  $\mathfrak{F}$  gibt es ein  $C < \infty$  mit  $|f_0| \leq C$ . Damit ist

$$|M(f_0, \tilde{g})| \leq M(f_{0+}, \tilde{g}_+) + M(f_{0+}, \tilde{g}_-) + M(f_{0-}, \tilde{g}_+) + M(f_{0-}, \tilde{g}_-) \\ \leq 2 \cdot M(C, \tilde{g}_+) + 2 \cdot M(C, \tilde{g}_-) = 2 \cdot C \cdot \|\tilde{g}\|_1$$

für alle  $g \in \mathcal{L}^1(\mathfrak{X}, \nu)$  und folglich  $M(f_0, \cdot)$  stetig. Da  $L^1(\mathfrak{X}, \nu) = L^\infty(\mathfrak{X}, \nu)$  ist, gibt es genau ein  $\bar{M}(f_0) \in L^\infty(\mathfrak{X}, \nu)$  mit  $M(f_0, \tilde{g}) = \int \bar{f}_0 \cdot g \, d\nu$  für alle  $g \in \mathcal{L}^1(\mathfrak{X}, \nu)$ , wobei  $\bar{f}_0 \in \bar{M}(f_0)$  ist.

Die letzte Gleichung zeigt, daß die Zuordnung  $f \mapsto \bar{M}(f)$  für alle  $f \in \mathfrak{F}$  eine positive, normierte lineare Abbildung von  $\mathfrak{F}$  nach  $L^\infty(\mathfrak{X}, \nu)$  liefert und daß die Zuordnung  $M \mapsto \bar{M}$  für alle positiven, normierten Bilinearformen  $M$  auf  $\mathfrak{F} \times L^1(\mathfrak{X}, \nu)$  den gewünschten Isomorphismus liefert.

Für das folgende Lemma benötigen wir eine Vorbemerkung.

$\varphi$  sei ein schwacher Markovkern von dem Maßraum  $(Y, \mathfrak{L}, \nu)$  nach einem Maßraum  $(X, \mathcal{O})$ .

Für eine  $\mathcal{O}$ -Elementarfunktion  $f$ , d.h. eine Funktion der Gestalt  $f = \sum_{i=1}^n a_i \cdot 1_{A_i}$  mit  $a_i \geq 0$  und  $A_i \in \mathcal{O}$  für  $i = 1, \dots, n$ , sei eine weitere Darstellung  $f = \sum_{j=1}^m b_j \cdot 1_{B_j}$  mit  $b_j \geq 0$  und  $B_j \in \mathcal{O}$  für  $j = 1, \dots, m$  gegeben. Für alle  $y$  aus dem Komplement einer  $\nu$ -Nullmenge aus  $\mathfrak{L}$  sind dann die Summen

$\sum_{i=1}^n a_i \cdot \varphi(A_i, y)$  und  $\sum_{j=1}^m b_j \cdot \varphi(B_j, y)$  definiert und endlich.

Weiter sind sie dort  $\nu$ -fast überall gleich, wie sich analog [2], S. 47 Lemma 10.2, ergibt. Unter  $\int f(x) \varphi(dx, \cdot)$  verstehen wir eine  $\mathfrak{L}$ -meßbare Funktion, die  $\nu$ -fast überall einer derartigen Summe gleich ist.

Sei  $f$  nun eine überall nicht negative  $\mathcal{O}$ -meßbare Funktion.

Nach [2], S. 53 Satz 11.6, gibt es eine isotone Folge

$(f_n)_{n \in \mathbb{N}}$  von  $\mathcal{O}$ -Elementarfunktionen mit  $\sup_{n \in \mathbb{N}} f_n = f$ .  $(g_n)_{n \in \mathbb{N}}$

sei eine weitere Folge mit dieser Eigenschaft. Dann gilt

$\sup_{n \in \mathbb{N}} \int f_n(x) \varphi(dx, \cdot) = \sup_{n \in \mathbb{N}} \int g_n(x) \varphi(dx, \cdot)$   $\nu$ -fast überall.

Dies ergibt sich analog [2], S. 50 Korollar 11.2. Unter

$\int f(x) \varphi(dx, \cdot)$  verstehen wir dann eine  $\mathfrak{L}$ -meßbare Funktion, die mit  $\sup_{n \in \mathbb{N}} \int f_n(x) \varphi(dx, \cdot)$   $\nu$ -fast überall übereinstimmt.

(10.4) Definition.  $f$  sei eine  $\mathcal{O}$ -meßbare Funktion derart,

daß  $\int f_+(x) \varphi(dx, \cdot)$  und  $\int f_-(x) \varphi(dx, \cdot)$  endlichwertig gewählt werden können. Dann sei  $\int f(x) \varphi(dx, \cdot)$  eine  $\mathfrak{L}$ -meßbare

Funktion, die  $\nu$ -fast überall gleich

$\int f_+(x) \varphi(dx, \cdot) - \int f_-(x) \varphi(dx, \cdot)$  ist.

Jetzt sei  $\mathfrak{F}$  wieder ein Vektorverband reeller Funktionen mit 1 auf  $X$ , und  $(Y, \mathfrak{L}, \nu)$  sei wie oben gegeben.

(10.5) Lemma.  $T: \mathcal{F} \rightarrow L^\infty(X, \nu)$  sei eine positive, normierte lineare Abbildung mit  $T(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} T(f_n)$  für jede isotone Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen aus  $\mathcal{F}_+$  mit  $\sup_{n \in \mathbb{N}} f_n \in \mathcal{F}$ . Dann gibt es einen schwachen Markovkern  $\varphi$  von  $(Y, \mathcal{Y}, \nu)$  nach  $(X, \mathcal{A}(\mathcal{F}))$  mit  $T(f) = \int f \varphi(dx, \cdot)$  für alle  $f \in \mathcal{F}$ .

Beweis.  $L^\infty(X, \nu)$  ist ordnungsvollständig, und es existiert sogar das Supremum jeder Familie von Äquivalenzklassen positiver  $\mathcal{X}$ -meßbarer Funktionen. Damit haben die Äquivalenzklassen  $\mathcal{X}$ -meßbarer Funktionen die Eigenschaften von  $\bar{\mathbb{R}}$ , die beim Beweis des Satzes von Daniell-Stone ([2], Satz 39.4) benötigt werden. Ersetzen wir also in dem Beweis  $\bar{\mathbb{R}}$  durch die Menge der Äquivalenzklassen  $\mathcal{X}$ -meßbarer Funktionen, so folgt unsere Behauptung. Wie die einzelnen Beweisschritte nach der Ersetzung von  $\bar{\mathbb{R}}$  durch die Menge der Äquivalenzklassen  $\mathcal{X}$ -meßbarer Funktionen sind, wollen wir jetzt angeben.

Zuerst definiert man für jede isotone Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen aus  $\mathcal{F}_+$   $T^*(\sup_{n \in \mathbb{N}} f_n) := \sup_{n \in \mathbb{N}} T(f_n)$ , wobei gezeigt wird, daß die Definition nur von  $\sup_{n \in \mathbb{N}} f_n$  abhängt.

Für jedes  $G$  aus

$$\mathcal{Q}(\mathcal{F}) := \{Q \in \mathcal{P}(X) : 1_Q = \sup_{n \in \mathbb{N}} f_n, (f_n)_{n \in \mathbb{N}} \text{ isoton}, f_n \in \mathcal{F}_+\}$$

setzt man dann

$$M^*(G) := T^*(1_G) \text{ und für alle } Q \in \mathcal{P}(X)$$

$$M^*(Q) := \inf\{M^*(G) : G \in \mathcal{Q}(\mathcal{F}), Q \leq G\}.$$

Das Infimum existiert, da  $L^\infty(X, \nu)$  ordnungsvollständig ist und da für alle  $G \in \mathcal{Q}(\mathcal{F})$   $\tilde{0} \leq M^*(G) \leq \tilde{1}$  wegen der Normiertheit und Positivität von  $T$  ist.

Für  $M^*$  gilt dann  $M^*(\emptyset) = \tilde{0}$ ,  $M^*(Q) \geq \tilde{0}$  für alle  $Q \in \mathcal{P}(X)$  und

$M^*(\bigcup_{n=1}^{\infty} Q_n) \leq \sum_{n=1}^{\infty} M^*(Q_n)$  für jede Folge  $(Q_n)_{n \in \mathbb{N}}$  von Teilmengen von  $X$ .

$\mathcal{A}^* := \{A \subset X : M^*(Q) \geq M^*(Q \cap A) + M^*(Q \setminus A) \text{ für alle } Q \in \mathcal{P}(X)\}$

ist eine  $\sigma$ -Algebra, die  $\mathcal{A}(\mathcal{F})$  umfaßt. Auf  $\mathcal{A}^*$  gilt für jede Folge paarweiser disjunkter Mengen  $(A_n)_{n \in \mathbb{N}}$

$$M^*(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} M^*(A_n).$$

Die Einschränkung von  $M^*$  auf  $\mathcal{A}(\mathcal{F})$  sei  $M$ . Jede Auswahl eines Repräsentantensystems von  $M(\mathcal{A}(\mathcal{F})) \subset L^\infty(Y, \nu)$  liefert einen schwachen Markovkern  $\varphi$  von  $(Y, \mathcal{Y}, \nu)$  nach  $(X, \mathcal{A}(\mathcal{F}))$  mit  $T(f) = \int \widehat{f \varphi(dx, \cdot)}^\nu$  für alle  $f \in \mathcal{F}$ .

Für den Fall, daß  $X$  ein kompakter topologischer Raum ist und  $\mathcal{F} = \mathcal{C}(X)$ , gilt das folgende

(10.6) Korollar.  $T: \mathcal{C}(X) \rightarrow L^\infty(Y, \nu)$  sei eine positive, normierte lineare Abbildung. Dann gibt es einen schwachen Markovkern  $\varphi$  von  $(Y, \mathcal{Y}, \nu)$  nach  $(X, \mathcal{A}(\mathcal{C}(X)))$  mit  $T(f) = \int \widehat{f \varphi(dx, \cdot)}^\nu$  für alle  $f \in \mathcal{C}(X)$ .

Beweis. Wegen der Positivität und Normiertheit von  $T$  gilt zunächst  $|T(f)| \leq \sup_{x \in X} |f(x)|$  für alle  $f \in \mathcal{C}(X)$ . Mit Hilfe des Satzes von Dini ([2], Korollar 43.2) folgt dann die in (10.5) geforderte Stetigkeitsbedingung.

## § 11. Entscheidungsfunktionen

Hier soll zunächst gezeigt werden, wie die Entscheidungsfunktionen im Sinne von Le Cam [11] und Heyer [8] sich in unseren Rahmen einfügen. Diese Autoren definieren als Experiment ein Tripel  $(X, E, (P_j)_{j \in \Theta})$ , wobei  $X$  eine nicht leere Menge,  $E$  ein Banachverband reeller beschränkter Funktionen mit 1 unter der Supremumsnorm und für alle  $j \in \Theta$   $P_j$  ein positives, normiertes lineares Funktional auf  $E$  ist. Die Familie  $(P_j)_{j \in \Theta}$  läßt sich also nicht ohne weiteres als Familie von normierten Maßen auf  $\mathcal{A}(E)$  deuten.

Als "Entscheidungsraum"  $D$  wird eine Teilmenge des  $\mathbb{R}^\Theta$  betrachtet und gefordert, daß  $D$  konvex und kompakt ist (in [11] noch dazu endliche lineare Dimension) und  $D \subset [-1, +1]^\Theta \subset \mathbb{R}^\Theta$  liegt. Als "Schaden" bei vorliegender Verteilung  $j$  und der Entscheidung  $t \in D$  ist dann  $\text{pr}_j(t)$  anzusehen.

### (A) Entscheidungsfunktionen nach Le Cam:

Wir fordern, was Le Cam nicht tut, zusätzlich, daß für alle  $j \in \Theta$   $P_j(\sup_{n \in \mathbb{N}} f_n) = \sup_{n \in \mathbb{N}} P_j(f_n)$  ist für jede isotone Folge  $(f_n)_{n \in \mathbb{N}}$  von Funktionen aus  $E_+$  mit  $\sup_{n \in \mathbb{N}} f_n \in E$ . Dann können wir die Familie  $(P_j)_{j \in \Theta}$  als Familie von normierten Maßen auf  $\mathcal{A}(E)$  auffassen. Wir können also das Experiment  $(X, \mathcal{A}(E), (P_j)_{j \in \Theta})$  betrachten und von dort zu einem zugehörigen schwach dominierten Experiment  $(X', \mathcal{A}(E)', (P'_j)_{j \in \Theta})$  übergehen (11.10). Das schwach dominierende Maß sei  $\nu$ . Als "Risikofunktionen" betrachtet Le Cam Abbildungen, die vermöge  $j \mapsto M(\text{pr}_j, \frac{\overline{P}_j}{d\nu})$  definiert sind, wobei  $M$  eine positive, normierte Bilinearform auf  $\mathcal{C}(D) \times L^1(\mathcal{A}(E)', \nu)$  ist.  $M$  stellt also in einem allgemeineren Sinne eine Entscheidungsfunktion

dar.

Nach (10.3) und (10.6) gibt es nun aber einen schwachen Markovkern  $\varphi$  von  $(X', \mathcal{A}(E)', \nu)$  nach  $(D, \mathcal{A}(\mathcal{C}(D)))$  derart, daß  $M(\text{pr}_{\mathcal{J}}, \frac{dP_{\mathcal{J}}}{d\nu}) = \int (\int \text{pr}_{\mathcal{J}}(t) \varphi(dt, \cdot)) \cdot \frac{dP_{\mathcal{J}}}{d\nu} d\nu$   
 $= \int (\int \text{pr}_{\mathcal{J}}(t) \varphi(dt, \cdot)) dP_{\mathcal{J}}'$  für alle  $\mathcal{J} \in \mathcal{C}$  ist. Wenn  $\varphi$  sogar ein Markovkern ist, ist also die Le Camsche Risikofunktion die übliche Risikofunktion der randomisierten Entscheidungsfunktion  $\varphi$  bei der Schadensfunktion  $\text{pr}_{\mathcal{J}}(t)$ ,  $\mathcal{J} \in \mathcal{C}$  und  $t \in D$ .

### (B) Entscheidungsfunktionen nach Heyer:

Heyer bezeichnet als Entscheidungsfunktionen bezüglich des Entscheidungsraums  $D$  alle Abbildungen  $f: X \rightarrow D$  mit  $f = (f_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$  so, daß  $f_{\mathcal{J}} \in E$  ist für jedes  $\mathcal{J} \in \mathcal{C}$ . In Proposition 3.1.2 zeigt er, daß die Menge der Entscheidungsfunktionen gleich der Menge aller  $(T(\text{pr}_{\mathcal{J}}))_{\mathcal{J} \in \mathcal{C}}$  ist, wobei  $T: \mathcal{C}(D) \rightarrow E$  eine positive, normierte lineare Abbildung ist. Da  $D$  kompakt ist, gibt es nun zu jedem derartigen  $T$  einen Markovkern  $\varphi$  von  $(X, \mathcal{A}(E))$  nach  $(D, \mathcal{A}(\mathcal{C}(D)))$  (10.2) mit  $T(\text{pr}_{\mathcal{J}}) = \int \text{pr}_{\mathcal{J}}(t) \varphi(dt, \cdot)$  für alle  $\mathcal{J} \in \mathcal{C}$ .  $\varphi$  ist aber eine randomisierte Entscheidungsfunktion im üblichen Sinne.

Für den Fall, daß  $E$  die Menge der meßbaren beschränkten Funktionen bezüglich einer  $\sigma$ -Algebra  $\mathcal{A}$  über  $X$  ist und  $(P_{\mathcal{J}})_{\mathcal{J} \in \mathcal{C}}$  eine durch ein dazu äquivalentes  $\sigma$ -endliches Maß  $\nu$  dominierte Familie von normierten Maßen auf  $\mathcal{A}$ , betrachtet Heyer als verallgemeinerte Entscheidungsfunktionen positive, normierte lineare Abbildungen  $T$  von  $\mathcal{C}(D)$  in  $L^{\infty}(\mathcal{A}, \nu)$ . Als Risikofunktion von einem solchen  $T$  wird die durch  $\mathcal{J} \mapsto P_{\mathcal{J}}(T(\text{pr}_{\mathcal{J}}))$  für alle  $\mathcal{J} \in \mathcal{C}$  definierte Abbildung betrachtet. Nach (10.6) gibt es dann aber einen schwachen Markovkern  $\varphi$  von  $(X, \mathcal{A}, \nu)$  nach  $(D, \mathcal{A}(\mathcal{C}(D)))$  mit  $P_{\mathcal{J}}(T(\text{pr}_{\mathcal{J}})) = \int (\int \text{pr}_{\mathcal{J}}(t) \varphi(dt, \cdot)) dP_{\mathcal{J}}$  für alle  $\mathcal{J} \in \mathcal{C}$ .

## § 12. Entscheidungstheoretische Charakterisierung der schwachen Suffizienz

$(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  und  $(Y, \mathcal{Y}, (Q_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$  seien im folgenden Experimente. In Anlehnung an die Überlegungen des vorigen Paragraphen treffen wir folgende Definitionen:

(12.1) Definition. Jede konvexe und kompakte Teilmenge  $D \subset [-1, +1]^{\Theta} \subset \mathbb{R}^{\Theta}$  heißt ein Entscheidungsraum für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  bzw.  $(Y, \mathcal{Y}, (Q_{\mathcal{Y}})_{\mathcal{Y} \in \Theta})$ .

(12.2) Definition. Jeder Markovkern  $\psi$  von  $(X, \mathcal{A})$  nach  $(D, \mathcal{A}(\mathcal{C}(D)))$ , wobei  $D$  ein beliebiger Entscheidungsraum für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  ist, heißt Entscheidungsfunktion für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$ .

(12.3) Definition.  $(P_{\mathcal{A}})_{\mathcal{A} \in \Theta}$  sei schwach dominiert durch das dazu äquivalente Maß  $\nu$ . Dann heißt jeder schwache Markovkern von  $(X, \mathcal{A}, \nu)$  nach  $(D, \mathcal{A}(\mathcal{C}(D)))$ , wobei  $D$  ein beliebiger Entscheidungsraum für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  ist, eine verallgemeinerte Entscheidungsfunktion für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$ .

(12.4) Definition. Sei  $D$  fest gegeben.

(a) Für jede Entscheidungsfunktion  $\psi$  für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  heißt die durch  $R_{\psi}(\mathcal{A}) := \int \left( \int_D \text{pr}_{\mathcal{A}}(t) \psi(dt, x) \right) P_{\mathcal{A}}(dx)$  für alle  $\mathcal{A} \in \Theta$  auf  $\Theta$  definierte Funktion  $R_{\psi}$  Risikofunktion der Entscheidungsfunktion  $\psi$ .

(b) Ist  $(P_{\mathcal{A}})_{\mathcal{A} \in \Theta}$  durch das dazu äquivalente Maß  $\nu$  schwach dominiert, so heißt für jede verallgemeinerte Entscheidungsfunktion  $\varphi$  für  $(X, \mathcal{A}, (P_{\mathcal{A}})_{\mathcal{A} \in \Theta})$  die durch  $R_{\varphi}(\mathcal{A}) := \int \left( \int_D \text{pr}_{\mathcal{A}}(t) \varphi(dt, x) \right) P_{\mathcal{A}}(dx)$  für alle  $\mathcal{A} \in \Theta$  auf  $\Theta$  definierte Funktion  $R_{\varphi}$  Risikofunktion der verallgemeinerten Entscheidungsfunktion  $\varphi$ .

Zufolge des vorangehenden Paragraphen befinden sich diese Definitionen mit denen von Heyer im Einklang.

(12.5) Satz. Die folgenden Aussagen sind äquivalent, wenn  $(X, \mathcal{A}, (P_\nu)_{\nu \in \mathcal{O}})$  durch ein zu  $(P_\nu)_{\nu \in \mathcal{O}}$  äquivalentes Maß schwach dominiert ist:

(1) Für jedes  $\varepsilon > 0$  gibt es für jeden Entscheidungsraum  $D$  und jede Entscheidungsfunktion  $\psi$  für  $(Y, \mathcal{Y}, (Q_\nu)_{\nu \in \mathcal{O}})$  eine verallgemeinerte Entscheidungsfunktion  $\varphi$  für  $(X, \mathcal{A}, (P_\nu)_{\nu \in \mathcal{O}})$  so, daß  $R_\varphi(\nu) \leq R_\psi(\nu) + \varepsilon$  für alle  $\nu \in \mathcal{O}$  ist.

(2) Es gibt einen schwachen Markovkern  $\chi$  von  $(X, \mathcal{A}, \nu)$  nach  $(Y, \mathcal{Y})$  mit  $Q_\nu(C) = \int \chi(C, \cdot) dP_\nu$  für alle  $\nu \in \mathcal{O}$  und  $C \in \mathcal{Y}$ .

Beweis. Nach [8], Korollar 4.2.3, ist die Existenz einer positiven, normierten linearen Abbildung  $T$  von den beschränkten  $\mathcal{Y}$ -meßbaren Funktionen in  $L^\infty(\mathcal{A}, \nu)$  mit  $T'P_\nu = Q_\nu$  für alle  $\nu \in \mathcal{O}$ , wobei  $T'$  die zu  $T$  adjungierte Abbildung sei, äquivalent mit (1). (In [8] ist  $\nu$  zwar  $\sigma$ -endlich vorausgesetzt, aber es wird nur die Beziehung  $L^{1'}(\mathcal{A}, \nu) = L^\infty(\mathcal{A}, \nu)$  gebraucht.) Wir zeigen, daß die Existenz der Abbildung  $T$  mit (2) äquivalent ist. Daß (2) die Existenz eines solchen  $T$  liefert, ist klar.

Existiere nun das  $T$ , und sei  $(f_n)_{n \in \mathbb{N}}$  eine isotone Folge überall nicht negativer Funktionen, die  $\mathcal{Y}$ -meßbar und beschränkt sind, mit  $\sup_{n \in \mathbb{N}} f_n < \infty$ . Für jede  $\mathcal{Y}$ -meßbare beschränkte Funktion  $f$  sei  $\hat{T}(f) \in T(f)$ . Gemäß der Definition der Adjungierten einer linearen Abbildung ist dann

$$\begin{aligned} \int \hat{T}(\sup_{n \in \mathbb{N}} f_n) dP_\nu &= \int \sup_{n \in \mathbb{N}} f_n dQ_\nu = \sup_{n \in \mathbb{N}} \int f_n dQ_\nu \\ &= \sup_{n \in \mathbb{N}} \int \hat{T}(f_n) dP_\nu = \int \sup_{n \in \mathbb{N}} \hat{T}(f_n) dP_\nu \text{ für alle } \nu \in \mathcal{O}. \end{aligned}$$

Außerdem gilt stets  $\sup_{n \in \mathbb{N}} \hat{T}(f_n) \leq \hat{T}(\sup_{n \in \mathbb{N}} f_n)$   $\nu$ -fast überall.

Zusammen folgt  $\sup_{n \in \mathbb{N}} T(f_n) = T(\sup_{n \in \mathbb{N}} f_n)$ .

Nach (10.5) gibt es dann einen schwachen Markovkern  $\chi$  von

$(X, \mathcal{A}, \nu)$  nach  $(Y, \mathcal{Z})$  mit  $T(1_C) = \widehat{\chi(C, \cdot)}$  für alle  $C \in \mathcal{X}$ .

Daher gilt  $Q_{\mathcal{V}}(C) = T'(P_{\mathcal{V}})(1_C) = P_{\mathcal{V}}(T(1_C)) = \int \chi(C, \cdot) dP_{\mathcal{V}}$   
für alle  $\mathcal{V} \in \mathcal{O}$  und  $C \in \mathcal{X}$ , was die Gültigkeit von (2) zeigt.

(12.6) Satz.  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  und  $(Y, \mathcal{Z}, (Q_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  seien schwach dominiert. Genau dann sind die beiden Experimente füreinander schwach suffizient, wenn es für jedes  $\varepsilon > 0$  und jeden Entscheidungsraum  $D$  zu jeder Entscheidungsfunktion  $\psi_1$  für  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  bzw.  $\psi_2$  für  $(Y, \mathcal{Z}, (Q_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  eine verallgemeinerte Entscheidungsfunktion  $\varphi_1$  für  $(Y, \mathcal{Z}, (Q_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  bzw.  $\varphi_2$  für  $(X, \mathcal{A}, (P_{\mathcal{V}})_{\mathcal{V} \in \mathcal{O}})$  gibt so, daß  $R_{\varphi_i}(\mathcal{V}) \leq R_{\psi_i}(\mathcal{V}) + \varepsilon$  für alle  $\mathcal{V} \in \mathcal{O}$  und  $i = 1, 2$  ist.

Beweis. Unter den angegebenen Voraussetzungen ist die schwache Suffizienz mit der Isomorphie äquivalent (7.2). Zweimalige Anwendung von (12.5) liefert dann die Behauptung.

Literaturverzeichnis

- [1] R. H. Bahadur, Sufficiency and statistical decision functions. *Ann. Math. Statist.* 25 (1954), 423-462.
- [2] H. Bauer, Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. Berlin (1968).
- [3] D. Blackwell, Comparison of experiments. *Proc. Second Berkeley Sympos. math. Statist. Prob.* (1951), 93-102.
- [4] D. Blackwell, Equivalent comparison of experiments. *Ann. Math. Statist.* 24 (1953), 265-272.
- [5] D. L. Burkholder, Sufficiency in the undominated case. *Ann. Math. Statist.* 32 (1961), 1191-1200.
- [6] I. Csizsár, Information-type measures of difference of probability distributions and indirect observations. *Stud. Sc. Math. Hung.* 2 (1967), 299-318.
- [7] P. R. Halmos and L. J. Savage, Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *Ann. Math. Statist.* 20 (1949), 225-241.
- [8] H. Heyer, Erschöpftheit und Invarianz beim Vergleich von Experimenten. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 12 (1969), 21-55.
- [9] A. Ionescu Tulcea, Topics in the theory of lifting. Berlin-Heidelberg-New York (1969).
- [10] S. Kakutani, Concrete representation of abstract L-spaces and the mean ergodic theorem. *Ann. Math.* 42 (1941), 523-537.
- [11] L. Le Cam, Sufficiency and approximate sufficiency. *Ann. Math. Statist.* 35 (1964), 1419-1455.
- [12] N. Morse and R. Sacksteder, Statistical isomorphism. *Ann. Math. Statist.* 37 (1966), 203-214.
- [13] T. S. Pitcher, Sets of measures not admitting necessary and sufficient statistics or subfields. *Ann. Math. Statist.* 28 (1957), 267-268.

- [14] R. Sacksteder, A note on statistical equivalence.  
Ann. Math. Statist. 38 (1967), 787-794.
- [15] A. C. Zaanen, Integration. Amsterdam (1967).

## Lebenslauf

Am 25. Oktober 1942 wurde ich als jüngstes von drei Kindern des Studienrates Hermann Mussmann und seiner Ehefrau Hanna geb. Dettmer in Nienburg an der Weser geboren.

Ostern 1949 kam ich in die Volksschule in Luthe bei Hannover. Von dort wechselte ich im Frühjahr 1953 auf die Mittelschule in der benachbarten Stadt Wunstorf über. Von Ostern 1956 an besuchte ich dann das dortige Gymnasium, an dem ich im Februar 1963 meine Reifeprüfung ablegte. Anschließend leistete ich bis zum Herbst 1964 meinen Grundwehrdienst bei der Bundeswehr ab.

Im Wintersemester 1964/65 begann ich ein Mathematikstudium an der Universität in München. Dort legte ich im Februar 1969 die Diplomprüfung ab. Mein Studium setzte ich danach fort, um zu promovieren.

Druck: Münchener Schnelldruck  
Leo Janikowski  
8 München 2  
Gabelsbergerstr. 19





